US005079352A

United States Patent [19]

Gelfand et al.

[11] Patent Number: 5,079,352

[45] Date of Patent: Jan. 7, 1992

[54] PURIFIED THERMOSTABLE ENZYME

[75] Inventors: **David H. Gelfand**, Oakland; **Susanne Stoffel**, El Cerrito; **Frances C. Lawyer**, Oakland; **Randall K. Saiki**, Richmond, all of Calif.

[73] Assignee: **Cetus Corporation**, Emeryville, Calif.

[21] Appl. No.: **523,394**

[22] Filed: **May 15, 1990**

Related U.S. Application Data

[63] Continuation of Ser. No. 143,441, Jan. 12, 1988, abandoned, which is a continuation of Ser. No. 63,509, Jun. 17, 1987, Pat. No. 4,889,818, which is a continuation of Ser. No. 899,241, Aug. 22, 1986, abandoned.

[51] **Int. Cl.**[5] **C12N 15/00**; C12N 9/00; C07H 15/12; C07H 17/00

[52] **U.S. Cl.** **536/27**; 435/172.3; 435/183; 435/320.1

[58] **Field of Search** 435/183, 194, 320.1, 435/172.3, 68, 70; 536/27; 935/12, 29

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,683,195	7/1987	Mullis et al.	435/6
4,683,202	7/1987	Mullis	435/91
4,788,135	11/1988	Davis et al.	435/6
4,889,818	12/1989	Gelfand et al.	435/194

FOREIGN PATENT DOCUMENTS

258017	3/1988	European Pat. Off. .
8906691	7/1989	PCT Int'l Appl. .

OTHER PUBLICATIONS

Kaledin et al., Nov. 1982, *Biochem* 47:1515–1521.
Kaledin et al., 1983, *Chem. Abst.* 98:298(49311q).
Ruttiman et al., 1985, *Eur. J. Biochem.* 149:41–46.
New England BioLabs Oct. 1987 Catalog Update.
Air et al., 1974, *FEBS Lett.* 38(3):277–281.
Fabry et al., 1976, *Biochim Biophys Acta.* 453(3):228–235.
Chien et al., 1976, *J. Bact.* 127(3):1550–1557.
Kaledin et al., 1980, *Biochem* 45(4):494–501.
Lawyer et al., 15 Apr. 1989, *J. Biol. Chem* 264(11):6427–6437.
Kaledin et al., 1981, *Biochem.* 46(9):1247–1254.
Sakaguchi and Yajima, 1974, *Fed. Proc.* 33(5):1492.
Chien et al., 1976, *Chem. Abst.* 85:180 (155559t).
Edgar et al., 1975, *ASM Abst.* 75:151(K26).
Kaledin et al., 1980, *Chem. Abst.* 93:377 (40169p).
Ruttiman et al., 1985, *Eur. J. Biochem.* 149:41–46.
Lucchini et al., 1985, *Curr. Genet.* 10:245–252.
Johnson et al., 1985, *Cell* 43(1):369–377.

Primary Examiner—Robert A. Wax
Assistant Examiner—Eric Steffe
Attorney, Agent, or Firm—Lisabeth F. Murphy; Kevin R. Kaster

[57] **ABSTRACT**

Recombinant DNA vectors that encode a thermostable DNA polymerase are useful in the recombinant production of thermostable DNA polymerase. The recombinant thermostable polymerase is preferred for use in the production of DNA in a polymerase chain reaction. Especially useful vectors encode the ~94,000 dalton thermostable DNA polymerase from *thermus aquaticus.*

27 Claims, 8 Drawing Sheets

EXHIBIT C

TAQ DNA POLYMERASE SEQUENCE

```
-120                -100                -80
      BglII                          PvuII
AAGCTCAGATCTACCTGCCTGAGGGCGTCCGGTTCCAGCTGGCCCTTCCCGAGGGGGAGA

-60                 -40                 -20
GGGAGGCGTTTCTAAAAGCCCTTCAGGACGCTACCCGGGGGCGGGTGGTGGAAGGGTAAC

1                   20                  40                  60
ATGAGGGGGATGCTGCCCCTCTTTGAGCCCAAGGGCCGGGTCCTCCTGGTGGACGGCCAC
MetArgGlyMetLeuProLeuPheGluProLysGlyArgValLeuLeuValAspGlyHis
  1

                    80                  100                 120
CACCTGGCCTACCGCACCTTCCACGCCCTGAAGGGCCTCACCACCAGCCGGGGGGAGCCG
HisLeuAlaTyrArgThrPheHisAlaLeuLysGlyLeuThrThrSerArgGlyGluPro

                   140                 160                 180
GTGCAGGCGGTCTACGGCTTCGCCAAGAGCCTCCTCAAGGCCCTCAAGGAGGACGGGGAC
ValGlnAlaValTyrGlyPheAlaLysSerLeuLeuLysAlaLeuLysGluAspGlyAsp
 41

                   200                 220                 240
GCGGTGATCGTGGTCTTTGACGCCAAGGCCCCCTCCTTCCGCCACGAGGCCTACGGGGGG
AlaValIleValValPheAspAlaLysAlaProSerPheArgHisGluAlaTyrGlyGly

                   260                 280                 300
TACAAGGCGGGCCGGGCCCCCACGCCGGAGGACTTTCCCCGGCAACTCGCCCTCATCAAG
TyrLysAlaGlyArgAlaProThrProGluAspPheProArgGlnLeuAlaLeuIleLys
 81

                   320                 340                 360
                              XhoI
GAGCTGGTGGACCTCCTGGGGCTGGCGCGCCTCGAGGTCCCGGGCTACGAGGCGGACGAC
GluLeuValAspLeuLeuGlyLeuAlaArgLeuGluValProGlyTyrGluAlaAspAsp
```

FIG.1-1

TAQ DNA POLYMERASE SEQUENCE

```
                380                 400                 420
GTCCTGGCCAGCCTGGCCAAGAAGGCGGAAAAGGAGGGCTACGAGGTCCGCATCCTCACC
ValLeuAlaSerLeuAlaLysLysAlaGluLysGluGlyTyrGluValArgIleLeuThr
121

                440                 460                 480
GCCGACAAAGACCTTTACCAGCTCCTTTCCGACCGCATCCACGTCCTCCACCCCGAGGGG
AlaAspLysAspLeuTyrGlnLeuLeuSerAspArgIleHisValLeuHisProGluGly

                500                 520                 540
Asp718
TACCTCATCACCCCGGCCTGGCTTTGGGAAAAGTACGGCCTGAGGCCCGACCAGTGGGCC
TyrLeuIleThrProAlaTrpLeuTrpGluLysTyrGlyLeuArgProAspGlnTrpAla
161

                560                 580                 600
GACTACCGGGCCCTGACCGGGGACGAGTCCGACAACCTTCCCGGGGTCAAGGGCATCGGG
AspTyrArgAlaLeuThrGlyAspGluSerAspAsnLeuProGlyValLysGlyIleGly

                620                 640                 660
              HindIII
GAGAAGACGGCGAGGAAGCTTCTGGAGGAGTGGGGGAGCCTGGAAGCCCTCCTCAAGAAC
GluLysThrAlaArgLysLeuLeuGluGluTrpGlySerLeuGluAlaLeuLeuLysAsn
201

                680                 700                 720
CTGGACCGGCTGAAGCCCGCCATCCGGGAGAAGATCCTGGCCCACATGGACGATCTGAAG
LeuAspArgLeuLysProAlaIleArgGluLysIleLeuAlaHisMetAspAspLeuLys

                740                 760                 780
CTCTCCTGGGACCTGGCCAAGGTGCGCACCGACCTGCCCCTGGAGGTGGACTTCGCCAAA
LeuSerTrpAspLeuAlaLysValArgThrAspLeuProLeuGluValAspPheAlaLys
241

                800                 820                 840
AGGCGGGAGCCCGACCGGGAGAGGCTTAGGGCCTTTCTGGAGAGGCTTGAGTTTGGCAGC
ArgArgGluProAspArgGluArgLeuArgAlaPheLeuGluArgLeuGluPheGlySer
```

FIG.1-2

TAQ DNA POLYMERASE SEQUENCE

```
                  860                 880                 900
                                    BstXI
CTCCTCCACGAGTTCGGCCTTCTGGAAAGCCCCAAGGCCCTGGAGGAGGCCCCCTGGCCC
LeuLeuHisGluPheGlyLeuLeuGluSerProLysAlaLeuGluGluAlaProTrpPro
281                           290

                  920                 940                 960
CCGCCGGAAGGGGCCTTCGTGGGCTTTGTGCTTTCCCGCAAGGAGCCCATGTGGGCCGAT
ProProGluGlyAlaPheValGlyPheValLeuSerArgLysGluProMetTrpAlaAsp

                  980                 1000                1020
CTTCTGGCCCTGGCCGCCGCCAGGGGGGGCCGGGTCCACCGGGCCCCCGAGCCTTATAAA
LeuLeuAlaLeuAlaAlaAlaArgGlyGlyArgValHisArgAlaProGluProTyrLys
321

                  1040                1060                1080
GCCCTCAGGGACCTGAAGGAGGCGCGGGGGCTTCTCGCCAAAGACCTGAGCGTTCTGGCC
AlaLeuArgAspLeuLysGluAlaArgGlyLeuLeuAlaLysAspLeuSerValLeuAla

                  1100                1120                1140
CTGAGGGAAGGCCTTGGCCTCCCGCCCGGCGACGACCCCATGCTCCTCGCCTACCTCCTG
LeuArgGluGlyLeuGlyLeuProProGlyAspAspProMetLeuLeuAlaTyrLeuLeu
361

                  1160                1180                1200
GACCCTTCCAACACCACCCCCGAGGGGGTGGCCCGGCGCTACGGCGGGGAGTGGACGGAG
AspProSerAsnThrThrProGluGlyValAlaArgArgTyrGlyGlyGluTrpThrGlu

                  1220                1240                1260
GAGGCGGGGGAGCGGGCCGCCCTTTCCGAGAGGCTCTTCGCCAACCTGTGGGGGAGGCTT
GluAlaGlyGluArgAlaAlaLeuSerGluArgLeuPheAlaAsnLeuTrpGlyArgLeu
401

                  1280                1300                1320
GAGGGGGAGGAGAGGCTCCTTTGGCTTTACCGGGAGGTGGAGAGGCCCCTTTCCGCTGTC
GluGlyGluGluArgLeuLeuTrpLeuTyrArgGluValGluArgProLeuSerAlaVal
```

FIG.1-3

TAQ DNA POLYMERASE SEQUENCE

```
                1340                1360                1380
CTGGCCCACATGGAGGCCACGGGGGTGCGCCTGGACGTGGCCTATCTCAGGGCCTTGTCC
LeuAlaHisMetGluAlaThrGlyValArgLeuAspValAlaTyrLeuArgAlaLeuSer
441

                1400        XhoI    1420                1440
CTGGAGGTGGCCGAGGAGATCGCCCGCCTCGAGGCCGAGGTCTTCCGCCTGGCCGGCCAC
LeuGluValAlaGluGluIleAlaArgLeuGluAlaGluValPheArgLeuAlaGlyHis

                1460        PvuII   1480                1500
CCCTTCAACCTCAACTCCCGGGACCAGCTGGAAAGGGTCCTCTTTGACGAGCTAGGGCTT
ProPheAsnLeuAsnSerArgAspGlnLeuGluArgValLeuPheAspGluLeuGlyLeu
481

                1520                1540                1560
CCCGCCATCGGCAAGACGGAGAAGACCGGCAAGCGCTCCACCAGCGCCGCCGTCCTGGAG
ProAlaIleGlyLysThrGluLysThrGlyLysArgSerThrSerAlaAlaValLeuGlu

                1580                1600        PstI    SacI    1620
GCCCTCCGCGAGGCCCACCCCATCGTGGAGAAGATCCTGCAGTACCGGGAGCTCACCAAG
AlaLeuArgGluAlaHisProIleValGluLysIleLeuGlnTyrArgGluLeuThrLys
521

                1640                1660                1680
CTGAAGAGCACCTACATTGACCCCTTGCCGGACCTCATCCACCCCAGGACGGGCCGCCTC
LeuLysSerThrTyrIleAspProLeuProAspLeuIleHisProArgThrGlyArgLeu

                1700                1720                1740
CACACCCGCTTCAACCAGACGGCCACGGCCACGGGCAGGCTAAGTAGCTCCGATCCCAAC
HisThrArgPheAsnGlnThrAlaThrAlaThrGlyArgLeuSerSerSerAspProAsn
561

                1760                1780        BamHI   1800
CTCCAGAACATCCCCGTCCGCACCCCGCTTGGGCAGAGGATCCGCCGGGCCTTCATCGCC
LeuGlnAsnIleProValArgThrProLeuGlyGlnArgIleArgArgAlaPheIleAla
```

FIG.1-4

TAQ DNA POLYMERASE SEQUENCE

```
                1820                1840                1860
                                    SacI
GAGGAGGGGTGGCTATTGGTGGCCCTGGACTATAGCCAGATAGAGCTCAGGGTGCTGGCC
GluGluGlyTrpLeuLeuValAlaLeuAspTyrSerGlnIleGluLeuArgValLeuAla
601

                1880                1900                1920
CACCTCTCCGGCGACGAGAACCTGATCCGGGTCTTCCAGGAGGGGCGGGACATCCACACG
HisLeuSerGlyAspGluAsnLeuIleArgValPheGlnGluGlyArgAspIleHisThr

                1940                1960                1980
        PvuII
GAGACCGCCAGCTGGATGTTCGGCGTCCCCCGGGAGGCCGTGGACCCCCTGATGCGCCGG
GluThrAlaSerTrpMetPheGlyValProArgGluAlaValAspProLeuMetArgArg
641

                2000                2020                2040
GCGGCCAAGACCATCAACTTCGGGGTCCTCTACGGCATGTCGGCCCACCGCCTCTCCCAG
AlaAlaLysThrIleAsnPheGlyValLeuTyrGlyMetSerAlaHisArgLeuSerGln

                2060                2080                2100
   NheI
GAGCTAGCCATCCCTTACGAGGAGGCCCAGGCCTTCATTGAGCGCTACTTTCAGAGCTTC
GluLeuAlaIleProTyrGluGluAlaGlnAlaPheIleGluArgTyrPheGlnSerPhe
681

                2120                2140                2160
CCCAAGGTGCGGGCCTGGATTGAGAAGACCCTGGAGGAGGGCAGGAGGCGGGGGTACGTG
ProLysValArgAlaTrpIleGluLysThrLeuGluGluGlyArgArgArgGlyTyrVal

                2180                2200                2220
GAGACCCTCTTCGGCCGCCGCCGCTACGTGCCAGACCTAGAGGCCCGGGTGAAGAGCGTG
GluThrLeuPheGlyArgArgArgTyrValProAspLeuGluAlaArgValLysSerVal
721
```

FIG.1-5

TAQ DNA POLYMERASE SEQUENCE

```
                2240                2260                2280
        .           .           .           .           .           .
CGGGAGGCGGCCGAGCGCATGGCCTTCAACATGCCCGTCCAGGGCACCGCCGCCGACCTC
ArgGluAlaAlaGluArgMetAlaPheAsnMetProValGlnGlyThrAlaAlaAspLeu
741

                2300                2320                2340
        .           .           .           .           .           .
ATGAAGCTGGCTATGGTGAAGCTCTTCCCCAGGCTGGAGGAAATGGGGGCCAGGATGCTC
MetLysLeuAlaMetValLysLeuPheProArgLeuGluGluMetGlyAlaArgMetLeu

                2360                2380                2400
                          XhoI
        .           .           .           .           .           .
CTTCAGGTCCACGACGAGCTGGTCCTCGAGGCCCCAAAAGAGAGGGCGGAGGCCGTGGCC
LeuGlnValHisAspGluLeuValLeuGluAlaProLysGluArgAlaGluAlaValAla
781

                2420                2440                2460
        .           .           .           .           .           .
CGGCTGGCCAAGGAGGTCATGGAGGGGGTGTATCCCCTGGCCGTGCCCCTGGAGGTGGAG
ArgLeuAlaLysGluValMetGluGlyValTyrProLeuAlaValProLeuGluValGlu

                2480                2500
        .           .           .           .
GTGGGGATAGGGGAGGACTGGCTCTCCGCCAAGGAGTGATACCACC
ValGlyIleGlyGluAspTrpLeuSerAlaLysGluEnd
821                                 832
```

FIG.1-6

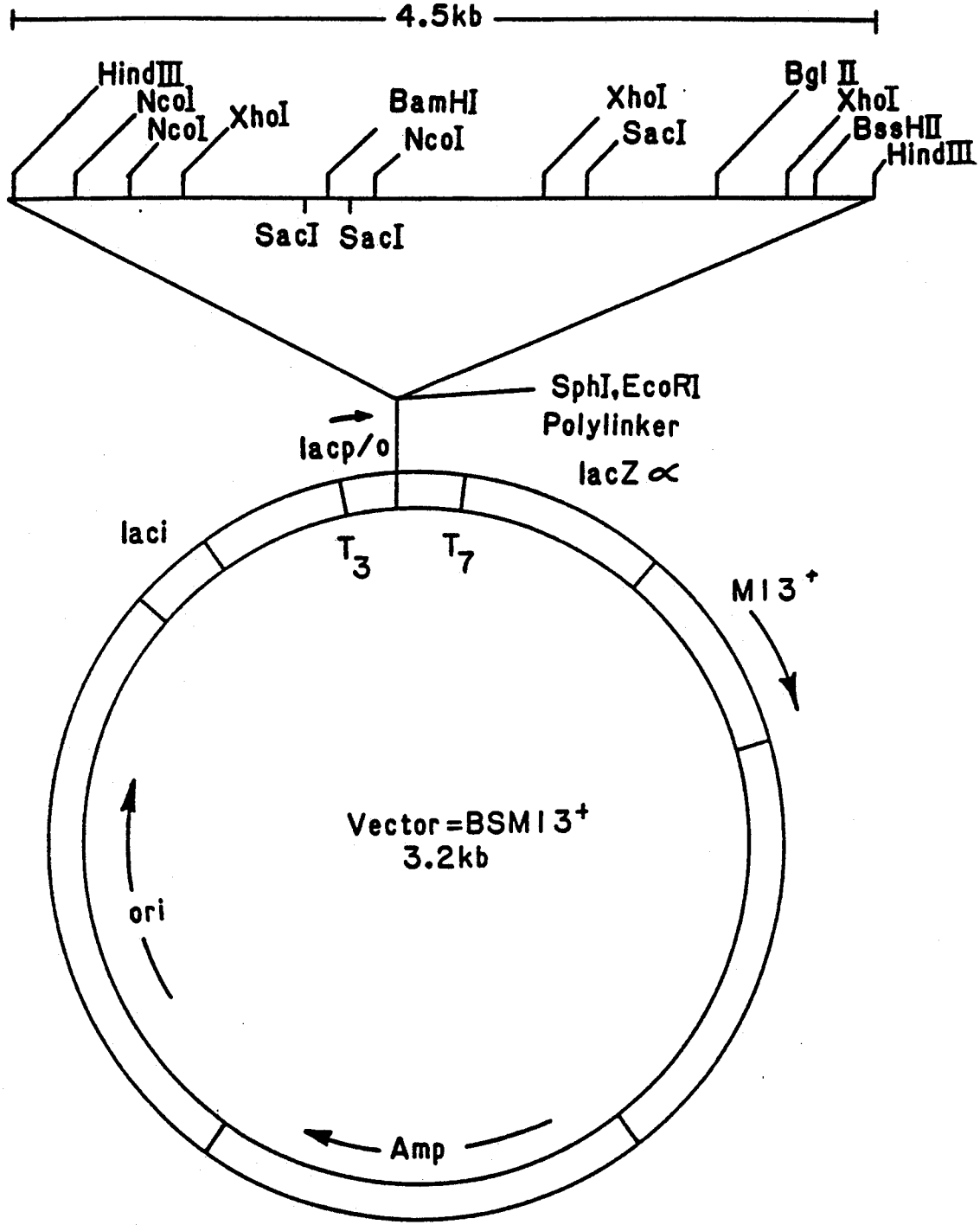

4.5kb
HindIII
NcoI
NcoI
XhoI
BamHI
NcoI
XhoI
SacI
Bgl II
XhoI
BssHII
HindIII
SacI
SacI
SphI,EcoRI
Polylinker
lacZ∝
lacp/o
laci
T3
T7
M13+
Vector=BSM13+
3.2kb
ori
Amp

FIG.2

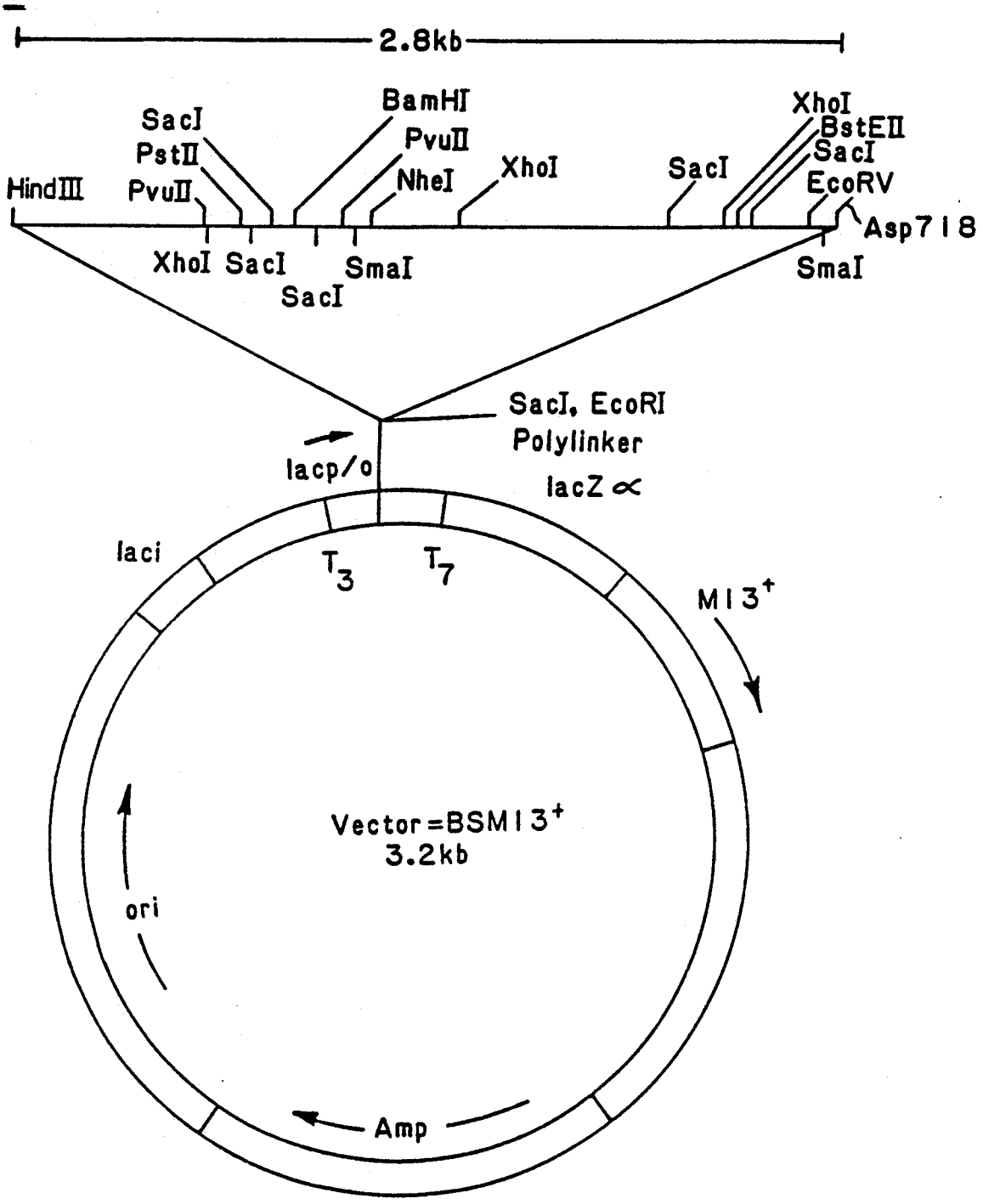

2.8kb
SacI
PstII
HindIII
PvuII
BamHI
PvuII
NheI
XhoI
XhoI
SacI
SacI
SmaI
SacI
XhoI
BstEII
SacI
EcoRV
Asp718
SmaI
SacI, EcoRI
Polylinker
lacZ∝
lacp/o
laci
T3
T7
M13+
Vector=BSM13+
3.2kb
ori
Amp

FIG.3

PURIFIED THERMOSTABLE ENZYME

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation of now abandoned Ser. No. 143,441, filed Jan. 12, 1988, which is a continuation in part of Ser. No. 063,509, filed June 17, 1987, which issued as U.S. Pat. No. 4,889,818, and which is a continuation in part of Ser. No. 899,241, filed Aug. 22, 1986, and now abandoned.

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to a purified thermostable enzyme. In one embodiment the enzyme is DNA polymerase purified from *Thermus aquaticus* and has a molecular weight of about 86,000–95,000. In another embodiment the enzyme is DNA polymerase produced by recombinant means.

2. Background Art

Extensive research has been conducted on the isolation of DNA polymerases from mesophilic microorganisms such as *E. coli*. See, for example, Bessman et al., *J. Biol. Chem.* (1957) 233: 171–177 and Buttin and Kornberg (1966) *J. Biol. Chem.* 241: 5419–5427.

In contrast, relatively little investigation has been made on the isolation and purification of DNA polymerases from thermophiles, such as *Thermus aquaticus*. Kaledin et al., *Biokymiya* (1980) 45: 644–651 discloses a six-step isolation and purification procedure of DNA polymerase from cells of *T. aquaticus* YT1 strain. These steps involve isolation of crude extract, DEAE-cellulose chromatography, fractionation on hydroxyapatite, fractionation on DEAE-cellulose, and chromatography on single-strand DNA-cellulose. The pools from each stage were not screened for contaminating endo- and exonuclease(s). The molecular weight of the purified enzyme is reported as 62,000 daltons per monomeric unit.

A second purification scheme for a polymerase from *T. aquaticus* is described by A. Chien et al., *J. Bacteriol.* (1976) 127: 1550–1557. In this process, the crude extract is applied to a DEAE-Sephadex column. The dialyzed pooled fractions are then subjected to treatment on a phosphocellulose column. The pooled fractions are dialyzed and bovine serum albumin (BSA) is added to prevent loss of polymerase activity. The resulting mixture is loaded on a DNA-cellulose column. The pooled material from the column is dialyzed and analyzed by gel filtration to have a molecular weight of about 63,000 daltons, and, by sucrose gradient centrifugation of about 68,000 daltons.

The use of a thermostable enzyme to amplify existing nucleic acid sequences in amounts that are large compared to the amount initially present has been suggested in U.S. Pat. No. 4,683,195. Primers, nucleotide triphosphates, and a polymerase are used in the process, which involves denaturation, synthesis of template strands and hybridization. The extension product of each primer becomes a template for the production of the desired nucleic acid sequence. The patent discloses that if the polymerase employed is a thermostable enzyme, it need not be added after every denaturation step, because the heat will not destroy its activity. No other advantages or details are provided on the use of a purified thermostable DNA polymerase. Furthermore, New England Biolabs had marketed a polymerase from *T. aquaticus*,

but was unaware that the polymeras activity decreased substantially with time in a storage buffer not containing non-ionic detergents.

Accordingly, there is a desire in the art to produce a purified, stable thermostable enzyme that may be used to improve the nucleic acid amplification process described above.

SUMMARY OF THE INVENTION

Accordingly, the present invention provides a purified thermostable enzyme that catalyzes combination of nucleotide triphosphates to form a nucleic acid strand complementary to a nucleic acid template strand. Preferably the purified enzyme is DNA polymerase from *Thermus aquaticus* and has a molecular weight of about 86,000–95,000 daltons. This purified material may be used in a temperature-cycling amplification reaction wherein nucleic acid sequences are produced from a given nucleic acid sequence in amounts that are large compared to the amount initially present so that they can be manipulated and/or analyzed easily.

The gene encoding the DNA polymerase enzyme from *Thermus aquaticus* has also been identified and cloned and provides yet another means to prepare the thermostable enzyme of the present invention. In addition to the gene encoding the approximately 86,000–95,000 dalton enzyme, gene derivatives encoding DNA polymerase activity are also presented.

The invention also encompasses a stable enzyme composition comprising a purified, thermostable enzyme as described above in a buffer containing one or more non-ionic polymeric detergents.

Finally, the invention provides a method of purification for the thermostable polymerase of the invention which comprises treating an aqueous mixture containing the thermostable polymerase with a hydrophobic interaction chromatographic support under conditions which promote hydrophobic interactions and eluting the bound thermostable polymerase from said support with a solvent which attenuates hydrophobic interactions.

The purified enzyme, as well as the enzymes produced by recombinant DNA techniques, provide much more specificity than the Klenow fragment, which is not thermostable, when used in the temperature-cycling amplification reaction. In addition, the purified enzyme and the recombinantly produced enzymes exhibit the appropriate activity expected when TTP or other nucleotide triphosphates are not present in the incubation mixture with the DNA template. Also, the enzymes herein have a broader pH profile than that of the thermostable enzyme from *Thermus aquaticus* described in the literature, with more than 50% of the activity at pH 6.4 as at pH 8.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is the DNA sequence and the predicted amino acid sequence for Taq polymerase. The amino acid sequence corresponding to the deduced primary translation product is numbered 1–832.

FIG. 2 is a restriction site map of plasmid pFC83 that contains the ~4.5 kb HindIII *T. aquaticus* DNA insert subcloned into plasmid BSM13+.

FIG. 3 is a restriction site map of plasmid pFC85 that contains the ~2.68 kb HindIII to Asp718 *T. aquaticus* DNA insert subcloned into plasmid BSM13+.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

As used herein, "cell", "cell line", and "cell culture" can be used interchangeably and all such designations include progeny. Thus, the words "transformants" or "transformed cells" includes the primary subject cell and cultures derived therefrom without regard for the number of transfers. It is also understood that all progeny may not be precisely identical in DNA content, due to deliberate or inadvertent mutations. Mutant progeny that have the same functionality as screened for in the originally transformed cell are included.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for procaryotes, for example, include a promoter, optionally an operator sequence, a ribosome binding site, and possibly, other as yet poorly understood sequences. Eucaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

The term "expression system" refers to DNA sequences containing a desired coding sequence and control sequences in operable linkage, so that hosts transformed with these sequences are capable of producing the encoded proteins. In order to effect transformation, the expression system may be included on a vector; however, the relevant DNA may then also be integrated into the host chromosome.

The term "gene" as used herein refers to a DNA sequence that encodes a recoverable bioactive polypeptide or precursor. The polypeptide can be encoded by a full-length gene sequence or any portion of the coding sequence so long as the enzymatic activity is retained.

In one embodiment of the invention, the DNA sequence encoding a full-length thermostable DNA polymerase of *Thermus aquaticus* (Taq) is provided. FIG. 1 shows this DNA sequence and the deduced amino acid sequence. For convenience, the amino acid sequence of this Taq polymerase will be used as a reference and other forms of the thermostable enzyme will be designated by referring to the sequence shown in FIG. 1. Since the N-terminal methionine may or may not be present, both forms are included in all cases wherein the thermostable enzyme is produced in bacteria.

"Operably linked" refers to juxtaposition such that the normal function of the components can be performed. Thus, a coding sequence "operably linked" to control sequences refers to a configuration wherein the coding sequences can be expressed under the control of the control sequences.

The term "mixture" as it relates to mixtures containing Taq polymerase refers to a collection of materials which includes Taq polymerase but which also includes alternative proteins. If the Taq polymerase is derived from recombinant host cells, the other proteins will ordinarily be those associated with the host. Where the host is bacterial, the contaminating proteins will, of course, be bacterial proteins.

"Non-ionic polymeric detergents" refers to surface-active agents that have no ionic charge and that are characterized, for purposes of this invention, by their ability to stabilize the enzyme herein at a pH range of from about 3.5 to about 9.5, preferably from 4 to 8.5.

The term "oligonucleotide" as used herein is defined as a molecule comprised of two or more deoxyribonucleotides or ribonucleotides, preferably more than three. Its exact size will depend on many factors, which in turn depend on the ultimate function or use of the oligonucleotide. The oligonucleotide may be derived synthetically or by cloning.

The term "primer" as used herein refers to an oligonucleotide, whether occurring naturally as in a purified restriction digest or produced synthetically, which is capable of acting as a point of initiation of synthesis when placed under conditions in which synthesis of a primer extension product which is complementary to a nucleic acid strand is initiated, i.e., in the presence of four different nucleotide triphosphates and thermostable enzyme in an appropriate buffer ("buffer" includes pH, ionic strength, cofactors, etc.) and at a suitable temperature. For Taq polymerase the buffer herein preferably contains 1.5–2 mM of a magnesium salt, preferably $MgCl_2$, 150–200 μM of each nucleotide, and 1 μM of each primer, along with preferably 50 mM KCl, 10 mM Tris buffer, pH 8–8.4, and 100 μg/ml gelatin.

The primer is preferably single-stranded for maximum efficiency in amplification, but may alternatively be double-stranded. If double-stranded, the primer is first treated to separate its strands before being used to prepare extension products. Preferably, the primer is an oligodeoxyribonucleotide. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the thermostable enzyme. The exact lengths of the primers will depend on many factors, including temperature, source of primer and use of the method. For example, depending on the complexity of the target sequence, the oligonucleotide primer typically contains 15–25 nucleotides, although it may contain more or fewer nucleotides. Short primer molecules generally require cooler temperatures to form sufficiently stable hybrid complexes with template.

The primers herein are selected to be "substantially" complementary to the different strands of each specific sequence to be amplified. This means that the primers must be sufficiently complementary to hybridize with their respective strands. Therefore, the primer sequence need not reflect the exact sequence of the template. For example, a non-complementary nucleotide fragment may be attached to the 5′ end of the primer, with the remainder of the primer sequence being complementary to the strand. Alternatively, non-complementary bases or longer sequences can be interspersed into the primer, provided that the primer sequence has sufficient complementarity with the sequence of the strand to be amplified to hybridize therewith and thereby form a template for synthesis of the extension product of the other primer. However, for detection purposes, particulary using labeled sequence-specific probes, the primers typically have exact complementarity to obtain the best results.

As used herein, the terms "restriction endonucleases" and "restriction enzymes" refer to bacterial enzymes each of which cut double-stranded DNA at or near a specific nucleotide sequence.

As used herein, the term "thermostable enzyme" refers to an enzyme which is stable to heat and is heat resistant and catalyzes (facilitates) combination of the nucleotides in the proper manner to form the primer extension products that are complementary to each nucleic acid strand. Generally, the synthesis will be initiated at the 3′ end of each primer and will proceed in the 5′ direction along the template strand, until synthesis terminates, producing molecules of different lengths.

There may be a thermostable enzyme, however, which initiates synthesis at the 5' end and proceeds in the other direction, using the same process as described above.

The thermostable enzyme herein must satisfy a single criterion to be effective for the amplification reaction, i.e., the enzyme must not become irreversibly denatured (inactivated) when subjected to the elevated temperatures for the time necessary to effect denaturation of double-stranded nucleic acids. Irreversible denaturation for purposes herein refers to permanent and complete loss of enzymatic activity. The heating conditions necessary for nucleic acid denaturation will depend, e.g., on the buffer salt concentration and composition and the length and nucleotide composition of the nucleic acids being denatured, but typically range from about 90° to about 105° C. for a time depending mainly on the temperature and the nucleic acid length, typically about 0.5 to four minutes. Higher temperatures may be tolerated as the buffer salt concentration and/or GC composition of the nucleic acid is increased. Preferably, the enzyme will not become irreversibly denatured at about 90°–100° C.

The thermostable enzyme herein preferably has an optimum temperature at which it functions that is higher than about 40° C., which is the temperature below which hybridization of primer to template is promoted, although, depending on (1) salt concentration and composition and (2) composition and length of primer, hybridization can occur at higher temperature (e.g., 45°–70° C.). The higher the temperature optimum for the enzyme, the greater the specificity and/or selectivity of the primer-directed extension process. However, enzymes that are active below 40° C., e.g., at 37° C., are also within the scope of this invention provided they are heat-stable. Preferably, the optimum temperature ranges from about 50° to 90° C., more preferably 60°–80° C.

The thermostable enzyme herein may be obtained from any source and may be a native or recombinant protein. Examples of enzymes that have been reported in the literature as being resistant to heat include heat-stable polymerases, such as, e.g., polymerases extracted from the thermophilic bacteria *Thermus flavus, Thermus ruber, Thermus thermophilus, Bacillus stearothermophilus* (which has a somewhat lower temperature optimum than the others listed), *Thermus aquaticus, Thermus lacteus, Thermus rubens*, and *Methanothermus fervidus*. In addition, thermostable polymerases isolated from the thermophilic archaebacteria include, for example, *Sulfolobus solfataricus, Sulfolobus acidocaldarius, Thermoplasma acidophilum, Methanobacterium thermoautotrophicum*, and *Desulfurococcus mobilis*.

The thermostable enzyme of the invention has the amino acid sequence presented in FIG. 1. In addition, any thermostable polymerase containing at least 50% homology to any contiguous stretch of nine or more amino acids presented therein is also intended to be within the scope of the invention. This homology can be determined using commercially available data banks such as the European Molecular Biology Laboratory (EMBL) or Genbank. Moreover, as new thermostable polymerases are identified, specific regions of homology between the newly identified sequences and the Taq polymerase sequence may be determined using, for example, the Sequence Analysis Software Package of the Genetics Computer Group of the University of Wisconsin. Specific regions of homology include the following sequences (numbered according to the numbering of amino acids in FIG. 1): residues 190–204, 262–270, 569–587, 718–732, 743–759, and 778–790.

The preferred thermostable enzyme herein is a DNA polymerase isolated from *Thermus aquaticus*. Various strains thereof are available from the American Type Culture Collection, Rockville, Md., and are described by T. D. Brock, *J. Bact.* (1969) 98: 289–297, and by T. Oshima, *Arch. Microbiol.* (1978) 117: 189–196. One of these preferred strains is strain YT-1.

For recovering the native protein the cells are grown using any suitable technique. One such technique is described by Kaledin et al., *Biokhimiya* (1980), supra, the disclosure of which is incorporated herein by reference. Briefly, the cells are grown on a medium, in one liter, of nitrilotriacetic acid (100 mg), tryptone (3 g), yeast extract (3 g), succinic acid (5 g), sodium sulfite (50 mg), riboflavin (1 mg), K_2HPO_4 (522 mg), $MgSO_4$ (480 mg), $CaCl_2$ (222 mg), NaCl (20 mg), and trace elements. The pH of the medium is adjusted to 8.0±0.2 with KOH. The yield is increased up to 20 grams of cells/liter if cultivated with vigorous aeration at a temperature of 70° C. Cells in the late logarithmic growth stage (determined by absorbance at 550 nm) are collected by centrifugation, washed with a buffer and stored frozen at −20° C.

In another method for growing the cells, described in Chien et al., *J. Bacteriol.* (1976), supra, the disclosure of which is incorporated herein by reference, a defined mineral salts medium containing 0.3% glutamic acid supplemented with 0.1 mg/l biotin, 0.1 mg/l thiamine, and 0.05 mg/l nicotinic acid is employed. The salts include nitrilotriacetic acid, $CaSO_4$, $MgSO_4$, NaCl, KNO_3, $NaNO_3$, $ZnSO_4$, H_3BO_3, $CuSO_4$, $NaMoO_4$, $CoCl_2$, $FeCl_3$, $MnSO_4$, and Na_2HPO_4. The pH of the medium is adjusted to 8.0 with NaOH.

In the Chien et al. technique, the cells are grown initially at 75° C. in a water bath shaker. On reaching a certain density, 1 liter of these cells is transferred to 16-liter carboys which are placed in hot-air incubators. Sterile air is bubbled through the cultures and the temperature maintained at 75° C. The cells are allowed to grow for 20 hours before being collected by centrifuge.

After cell growth, the isolation and purification of the enzyme take place in six stages, each of which is carried out at a temperature below room temperature, preferably about 4° C.

In the first stage or step, the cells, if frozen, are thawed, disintegrated by ultrasound, suspended in a buffer at about pH 7.5, and centrifuged.

In the second stage, the supernatant is collected and then fractionated by adding a salt such as dry ammonium sulfate. The appropriate fraction (typically 45–75% of saturation) is collected, dissolved in a 0.2M potassium phosphate buffer preferably at pH 6.5, and dialyzed against the same buffer.

The third step removes nucleic acids and some protein. The fraction from the second stage is applied to a DEAE-cellulose column equilibrated with the same buffer as used above. Then the column is washed with the same buffer and the flow-through protein-containing fractions, determined by absorbance at 280 nm, are collected and dialyzed against a 10 mM potassium phosphate buffer, preferably with the same ingredients as the first buffer, but at a pH of 7.5.

In the fourth step, the fraction so collected is applied to a hydroxyapatite column equilibrated with the buffer used for dialysis in the third step. The column is then washed and the enzyme eluted with a linear gradient of

a buffer such as 0.01M to 0.5M potassium phosphate buffer at pH 7.5 containing 10 mM 2-mercaptoethanol and 5% glycerine. The pooled fractions containing thermostable enzyme (e.g., DNA polymerase) activity are dialyzed against the same buffer used for dialysis in the third step.

In the fifth stage, the dialyzed fraction is applied to a DEAE-cellulose column, equilibrated with the buffer used for dialysis in the third step. The column is then washed and the enzyme eluted with a linear gradient of a buffer such as 0.01 to 0.6M KCl in the buffer used for dialysis in the third step. Fractions with thermostable enzyme activity are then tested for contaminating deoxyribonucleases (endo- and exonucleases) using any suitable procedure. For example, the endonuclease activity may be determined electrophoretically from the change in molecular weight of phage λ DNA or supercoiled plasmid DNA after incubation with an excess of DNA polymerase. Similarly, exonuclease activity may be determined electrophoretically from the change in molecular weight of DNA after treatment with a restriction enzyme that cleaves at several sites.

The fractions determined to have no deoxyribonuclease activity are pooled and dialyzed against the same buffer used in the third step.

In the sixth step, the pooled fractions are placed on a phosphocellulose column with a set bed volume. The column is washed and the enzyme eluted with a linear gradient of a buffer such as 0.01 to 0.4M KCl in a potassium phosphate buffer at pH 7.5. The pooled fractions having thermostable polymerase activity and no deoxyribonuclease activity are dialyzed against a buffer at pH 8.0.

The molecular weight of the dialyzed product may be determined by any technique, for example, by SDS-PAGE analysis using protein molecular weight markers. The molecular weight of one of the preferred enzymes herein, the DNA polymerase purified from *Thermus aquaticus*, is determined by the above method to be about 86,000–90,000 daltons. The molecular weight of this same DNA polymerase as determined by the predicted amino acid sequence is calculated to be approximately 94,000 daltons. Thus, the molecular weight of the full length DNA polymerase is dependent upon the method employed to determine this number and falls within the range of 86,000–95,000 daltons.

The thermostable enzyme of this invention may also be produced by recombinant DNA techniques, as the gene encoding this enzyme has been cloned from *Thermus aquaticus* genomic DNA. The complete coding sequence for the *Thermus aquaticus* (Taq) polymerase can be derived from bacteriophage CH35: Taq#4-2 on an approximately 3.5 kilobase (kb) BglII-Asp718 (partial) restriction fragment contained within an ~18 kb genomic DNA insert fragment. This bacteriophage was deposited with the American Type Culture Collection (ATCC) on May 29, 1987 and has accession no. 40,336. Alternatively, the gene can be constructed by ligating an ~730 base pair (bp) BglII-HindIII restriction fragment isolated from plasmid pFC83 (ATCC 67,422 deposited May 29, 1987) to an ~2.68 kb HindIII-Asp718 restriction fragment isolated from plasmid pFC85 (ATCC 67,421 deposited May 29, 1987). The pFC83 restriction fragment comprises the amino-terminus of the Taq polymerase gene while the restriction fragment from pFC85 comprises the carboxy-terminus. Thus, ligation of these two fragments into a correspondingly digested vector with appropriate control sequences will result in the translation of a full-length Taq polymerase.

As stated previously, the DNA and deduced amino acid sequence of a preferred thermostable enzyme is provided in FIG. 1. In addition to the N-terminal deletion described supra, it has also been found that the entire coding sequence of the Taq polymerase gene is not required to recover a biologically active gene product with DNA polymerase activity. Amino-terminal deletions wherein approximately one-third of the coding sequence is absent has resulted in producing a gene product that is quite active in polymerase assays.

In addition to the N-terminal deletions, individual amino acid residues in the peptide chain comprising Taq polymerase may be modified by oxidation, reduction, or other derivatization, and the protein may be cleaved to obtain fragments that retain activity. Such alterations that do not destroy activity do not remove the protein from the definition, and are specifically included.

Thus, modifications to the primary structure itself by deletion, addition, or alteration of the amino acids incorporated into the sequence during translation can be made without destroying the high temperature DNA polymerase activity of the protein. Such substitutions or other alterations result in proteins having an amino acid sequence encoded by DNA falling within the contemplated scope of the present invention.

Polyclonal antiserum from rabbits immunized with the purified 86,000–95,000 dalton polymerase of this invention was used to probe a *Thermus aquaticus* partial genomic expression library to obtain the appropriate coding sequence as described below. The cloned genomic sequence can be expressed as a fusion polypeptide, expressed directly using its own control sequences, or expressed by constructions using control sequences appropriate to the particular host used for expression of the enzyme.

Of course, the availability of DNA encoding these sequences provides the opportunity to modify the codon sequence so as to generate mutein (mutant protein) forms also having DNA polymerase activity.

Thus, these tools can provide the complete coding sequence for Taq DNA polymerase from which expression vectors applicable to a variety of host systems can be constructed and the coding sequence expressed. Portions of the Taq polymerase-encoding sequence are useful as probes to retrieve other thermostable polymerase-encoding sequences in a variety of species. Accordingly, portions of the genomic DNA encoding at least four to six amino acids can be replicated in *E. coli* and the denatured forms used as probes or oligodeoxyribonucleotide probes can be synthesized which encode at least four to six amino acids and used to retrieve additional DNAs encoding a thermostable polymerase. Because there may not be a precisely exact match between the nucleotide sequence in the *Thermus aquaticus* form and that in the corresponding portion of other species, oligomers containing approximately 12–18 nucleotides (encoding the four to six amino acid stretch) are probably necessary to obtain hybridization under conditions of sufficient stringency to eliminate false positives. The sequences encoding six amino acids would supply information sufficient for such probes.

SUITABLE HOSTS, CONTROL SYSTEMS AND METHODS

In general terms, the production of a recombinant form of Taq polymerase typically involves the following:

First, a DNA is obtained that encodes the mature (used here to include all muteins) enzyme or a fusion of the Taq polymerase to an additional sequence that does not destroy its activity or to an additional sequence cleavable under controlled conditions (such as treatment with peptidase) to give an active protein. If the sequence is uninterrupted by introns it is suitable for expression in any host. This sequence should be in an excisable and recoverable form.

The excised or recovered coding sequence is then preferably placed in operable linkage with suitable control sequences in a replicable expression vector. The vector is used to transform a suitable host and the transformed host cultured under favorable conditions to effect the production of the recombinant Taq polymerase. Optionally the Taq polymerase is isolated from the medium or from the cells; recovery and purification of the protein may not be necessary in some instances, where some impurities may be tolerated.

Each of the foregoing steps can be done in a variety of ways. For example, the desired coding sequences may be obtained from genomic fragments and used directly in appropriate hosts. The constructions for expression vectors operable in a variety of hosts are made using appropriate replicons and control sequences, as set forth below. Suitable restriction sites can, if not normally available, be added to the ends of the coding sequence so as to provide an excisable gene to insert into these vectors.

The control sequences, expression vectors, and transformation methods are dependent on the type of host cell used to express the gene. Generally, procaryotic, yeast, insect or mammalian cells are presently useful as hosts. Procaryotic hosts are in general the most efficient and convenient for the production of recombinant proteins and therefore preferred for the expression of Taq polymerase.

In the particular case of Taq polymerase, evidence indicates that considerable deletion at the N-terminus of the protein may occur under both recombinant and native conditions, and that the DNA polymerase activity of the protein is still retained. It appears that the native proteins previously isolated may be the result of proteolytic degradation, and not translation of a truncated gene. The mutein produced from the truncated gene of plasmid pFC85 is, however, fully active in assays for DNA polymerase, as is that produced from DNA encoding the full-length sequence. Since it is clear that certain N-terminal shortened forms of the polymerase are active, the gene constructs used for expression of these polymerases may also include the corresponding shortened forms of the coding sequence.

CONTROL SEQUENCES AND CORRESPONDING HOSTS

Procaryotes most frequently are represented by various strains of *E. coli*. However, other microbial strains may also be used, such as bacilli, for example, *Bacillus subtilis*, various species of Pseudomonas, or other bacterial strains. In such procaryotic systems, plasmid vectors that contain replication sites and control sequences derived from a species compatible with the host are used. For example, *E. coli* is typically transformed using derivatives of pBR322, a plasmid derived from an *E. coli* species by Bolivar, et al., *Gene* (1977) 2: 95. pBR322 contains genes for amplicillin and tetracycline resistance, and thus provides additional markers that can be either retained or destroyed in constructing the desired vector. Commonly used procaryotic control sequences, which are defined herein to include promoters for transcription initiation, optionally with an operator, along with ribosome binding site sequences, include such commonly used promoters as the β-lactamase (penicillinase) and lactose (lac) promoter systems (Chang, et al., *Nature* (1977) 198: 1056), the tryptophan (trp) promoter system (Goeddel, et al., *Nucleic Acids Res.* (1980) 8: 4057) and the lambda-derived P_L promoter (Shimatake, et al., *Nature* (1981) 292: 128) and N-gene ribosome binding site, which has been made useful as a portable control cassette (as set forth in U.S. Pat. No. 4,711,845, issued Dec. 8, 1987), which comprises a first DNA sequence that is the P_L promoter operably linked to a second DNA sequence corresponding to N_{RBS} upstream of a third DNA sequence having at least one restriction site that permits cleavage within six bp 3' of the N_{RBS} sequence. Also useful is the phosphatase A (phoA) system described by Chang, et al. in European Patent Publication No. 196,864 published Oct. 8, 1986, assigned to the same assignee and incorporated herein by reference. However, any available promoter system compatible with procaryotes can be used.

In addition to bacteria, eucaryotic microbes, such as yeast, may also be used as hosts. Laboratory strains of *Saccharomyces cerevisiae*, Baker's yeast, are most used, although a number of other strains are commonly available. While vectors employing the 2 micron origin of replication are illustrated (Broach, J. R., *Meth. Enz.* (1983) 101: 307), other plasmid vectors suitable for yeast expression are known (see, for example, Stinchcomb, et al., *Nature* (1979) 282: 39, Tschempe, et al., *Gene* (1980) 10: 157 and Clarke, L., et al., *Meth. Enz.* (1983) 101: 300). Control sequences for yeast vectors include promoters for the synthesis of glycolytic enzymes (Hess, et al., *J. Adv. Enzyme Reg.* (1968) 7: 149; Holland, et al., *Biotechnology* (1978) 17: 4900).

Additional promoters known in the art include the promoter for 3-phosphoglycerate kinase (Hitzeman, et al., *J. Biol. Chem.* (1980) 255: 2073), and those for other glycolytic enzymes, such as glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase. Other promoters that have the additional advantage of transcription controlled by growth conditions are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, and enzymes responsible for maltose and galactose ultilization (Holland, supra).

It is also believed that terminator sequences are desirable at the 3' end of the coding sequences. Such terminators are found in the 3' untranslated region following the coding sequences in yeast-derived genes. Many of the vectors illustrated contain control sequences derived from the enolase gene containing plasmid peno46 (Holland, M. J., et al., *J. Biol. Chem.* (1981) 256: 1385) or the LEU2 gene obtained from YEp13 (Broach, J., et al., *Gene* (1978) 8: 121); however, any vector containing a

yeast-compatible promoter, origin of replication, and other control sequences is suitable.

It is also, of course, possible to express genes encoding polypeptides in eucaryotic host cell cultures derived from multicellular organisms. See, for example, *Tissue Culture*, Academic Press, Cruz and Patterson, editors (1973). Useful host cell lines include murine myelomas N51, VERO and HeLa cells, and Chinese hamster ovary (CHO) cells. Expression vectors for such cells ordinarily include promoters and control sequences compatible with mammalian cells such as, for example, the commonly used early and late promoters from Simian Virus 40 (SV 40) (Fiers, et al., *Nature* (1978) 273: 113), or other viral promoters such as those derived from polyoma, Adenovirus 2, bovine papiloma virus, or avian sarcoma viruses, or immunoglobulin promoters and heat shock promoters. A system for expressing DNA in mammalian systems using the BPV as a vector is disclosed in U.S. Pat. No. 4,419,446. A modification of this system is described in U.S. Pat. No. 4,601,978. General aspects of mammalian cell host system transformations have been described by Axel, U.S. Pat. No. 4,399,216. It now appears, also, that "enhancer" regions are important in optimizing expression; these are, generally, sequences found upstream of the promoter region. Origins of replication may be obtained, if needed, from viral sources. However, integration into the chromosome is a common mechanism for DNA replication in eucaryotes.

Plant cells are also now available as hosts, and control sequences compatible with plant cells such as the nopaline synthase promoter and polyadenylation signal sequences (Depicker, A., et al., *J. Mol. Appl. Gen.* (1982) 1: 561) are available.

Recently, in addition, expression systems employing insect cells utilizing the control systems provided by baculovirus vectors have been described (Miller, D. W., et al., in *Genetic Engineering* (1986) Setlow, J. K. et al., eds., Plenum Publishing, Vol. 8, pp. 277–297). These systems are also successful in producing Taq polymerase.

TRANSFORMATIONS

Depending on the host cell used, transformation is done using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described by Cohen, S. N., *Proc. Natl. Acad. Sci.* (U.S.A.) (1972) 69: 2110 is used for procaryotes or other cells that contain substantial cell wall barriers. Infection with *Agrobacterium tumefaciens* (Shaw, C. H., et al., *Gene* (1983) 23: 315) is used for certain plant cells. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, *Virology* (1978) 52: 546 is preferred. Transformations into yeast are carried out according to the method of Van Solingen, P., et al., *J. Bact.* (1977) 130: 946 and Hsiao, C. L., et al., *Proc. Natl. Acad. Sci.* (U.S.A.) (1979) 76: 3829.

CONSTRUCTION OF A λGT11 EXPRESSION LIBRARY

The strategy for isolating DNA encoding desired proteins, such as the Taq polymerase encoding DNA, using the bacteriophage vector lambda gt11, is as follows. A library can be constructed of EcoRI-flanked AluI fragments, generated by complete digestion of *Thermus aquaticus* DNA, inserted at the EcoRI site in the lambda gt11 phage (Young and Davis, *Proc. Natl. Acad. Sci* U.S.A. (1983) 80: 1194–1198). Because the unique EcoRI site in this bacteriophage is located in the carboxy-terminus of the β-galactosidase gene, inserted DNA (in the appropriate frame and orientation) is expressed as protein fused with β-galactosidase under the control of the lactose operon promoter/operator.

Genomic expression libraries are then screened using the antibody plaque hybridization procedure. A modification of this procedure, referred to as "epitope selection," uses antiserum against the fusion protein sequence encoded by the phage, to confirm the identification of hybridized plaques. Thus, this library of recombinant phages could be screened with antibodies that recognize the 86,000–95,000 dalton Taq polymerase in order to identify phage that carry DNA segments encoding the antigenic determinants of this protein.

Approximately 2×10^5 recombinant phage are screened using total rabbit Taq polymerase antiserum. In this primary screen, positive signals are detected and one or more of these phages are purified from candidate plaques which failed to react with preimmune serum and reacted with immune serum and analyzed in some detail. To examine the fusion proteins produced by the recombinant phage, lysogens of the phage in the host Y1089 are produced. Upon induction of the lysogens and gel electrophoresis of the resulting proteins, each lysogen may be observed to produce a new protein, not found in the other lysogens, or duplicate sequences may result. Phage containing positive signals are picked; in this case, one positive plaque was picked for further identification and replated at lower densities to purify recombinants and the purified clones were analyzed by size class via digestion with EcoRI restriction enzyme. Probes can then be made of the isolated DNA insert sequences and labeled appropriately and these probes can be used in conventional colony or plaque hybridization assays described in Maniatis et al., *Molecular Cloning: A Laboratory Manual* (1982), the disclosure of which is incorporated herein by reference.

The labeled probe was used to probe a second genomic library constructed in a Charon 35 bacteriophage (Wilhelmine, A. M. et al., *Gene* (1983) 26: 171–179). This library was made from Sau3A partial digestions of genomic *Thermus aquaticus* DNA and size fractionated fragments (15–20 kb) were cloned into the BamHI site of the Charon 35 phage. The probe was used to isolate phage containing DNA encoding the Taq polymerase. One of the resulting phage, designated CH35:Taq#4-2, was found to contain the entire gene sequence. Partial sequences encoding portions of the gene were also isolated.

VECTOR CONSTRUCTION

Construction of suitable vectors containing the desired coding and control sequences employs standard ligation and restriction techniques that are well understood in the art. Isolated plasmids, DNA sequences, or synthesized oligonucleotides are cleaved, tailored, and religated in the form desired.

Site-specific DNA cleavage is performed by treating with the suitable restriction enzyme (or enzymes) under conditions that are generally understood in the art, and the particulars of which are specified by the manufacturer of these commercially available restriction enzymes. See, e.g., New England Biolabs, Product Catalog. In general, about 1 μg of plasmid or DNA sequence is cleaved by one unit of enzyme in about 20 μl of buffer solution; in the examples herein, typically an excess of

restriction enzyme is used to ensure complete digestion of the DNA substrate. Incubation times of about one hour to two hours at about 37° C. are workable, although variations can be tolerated. After each incubation, protein is removed by extraction with phenol/-chloroform, and may be followed by ether extraction, and the nucleic acid recovered from aqueous fractions by precipitation with ethanol. If desired, size separation of the cleaved fragments may be performed by polyacrylamide gel or agarose gel electrophoresis using standard techniques. A general description of size separations is found in *Methods in Enzymology* (1980) 65: 499–560.

Restriction-cleaved fragments may be blunt-ended by treating with the large fragment of *E. coli* DNA polymerase I (Klenow) in the presence of the four deoxynucleotide triphosphates (dNTPs) using incubation times of about 15 to 25 minutes at 20° to 25° C. in 50 mM Tris pH 7.6, 50 mM NaCl, 10 mM $MgCl_2$, 10 mM DTT and 50–100 μM dNTPs. The Klenow fragment fills in at 5′ sticky ends, but chews back protruding 3′ single strands, even though the four dNTPs are present. If desired, selective repair can be performed by supplying only one of the, or selected, dNTPs within the limitations dictated by the nature of the sticky ends. After treatment with Klenow, the mixture is extracted with phenol/-chloroform and ethanol precipitated. Treatment under appropriate conditions with S1 nuclease results in hydrolysis of any single-stranded portion.

Synthetic oligonucleotides may be prepared using the triester method of Matteucci, et al., (*J. Am. Chem. Soc.* (1981) 103: 3185–3191) or using automated synthesis methods. Kinasing of single strands prior to annealing or for labeling is achieved using an excess, e.g., approximately 10 units of polynucleotide kinase to 1 nM substrate in the presence of 50 mM Tris, pH 7.6, 10 mM $MgCl_2$, 5 mM dithiothreitol, 1–2 mM ATP. If kinasing is for labeling of probe, the ATP will contain high specific activity γ-^{32}p.

Ligations are performed in 15–30 μl volumes under the following standard conditions and temperatures: 20 mM Tris-Cl pH 7.5, 10 mM $MgCl_2$, 10 mM DTT, 33 μg/ml BSA, 10 mM-50 mM NaCl, and either 40 μM ATP, 0.01–0.02 (Weiss) units T4 DNA ligase at 0° C. (for "sticky end" ligation) or 1 mM ATP, 0.3–0.6 (Weiss) units T4 DNA ligase at 14° C. (for "blunt end" ligation). Intermolecular "sticky end" ligations are usually performed at 33–100 μg/ml total DNA concentrations (5–100 nM total end concentration). Intermolecular blunt end ligations (usually employing a 10–30 fold molar excess of linkers) are performed at 1 μM total ends concentration.

In vector construction employing "vector fragments", the vector fragment is commonly treated with bacterial alkaline phosphatase (BAP) in order to remove the 5′ phosphate and prevent religation of the vector. BAP digestions are conducted at pH 8 in approximately 150 mM Tris, in the presence of Na^+ and Mg^{+2} using about 1 unit of BAP per mg of vector at 60° C. for about one hour. In order to recover the nucleic acid fragments, the preparation is extracted with phenol/chloroform and ethanol precipitated. Alternatively, religation can be prevented in vectors that have been double digested by additional restriction enzyme digestion of the unwanted fragments.

MODIFICATION OF DNA SEQUENCES

For portions of vectors derived from cDNA or genomic DNA that require sequence modifications, site-specific primer-directed mutagenesis is used. This technique is now standard in the art, and is conducted using a synthetic oligonucleotide primer complementary to a single-stranded phage DNA to be mutagenized except for limited mismatching, representing the desired mutation. Briefly, the synthetic oligonucleotide is used as a primer to direct synthesis of a strand complementary to the phage, and the resulting double-stranded DNA is transformed into a phage-supporting host bacterium. Cultures of the transformed bacteria are plated in top agar, permitting plaque formation from single cells that harbor the phage.

Theoretically, 50% of the new plaques will contain the phage having, as a single strand, the mutated form; 50% will have the original sequence. The plaques are transferred to nitrocellulose filters and the "lifts" hybridized with kinased synthetic primer at a temperature that permits hybridization of an exact match, but at which the mismatches with the original strand are sufficient to prevent hybridization. Plaques that hybridize with the probe are then picked and cultured, and the DNA is recovered.

VERIFICATION OF CONSTRUCTION

In the constructions set forth below, correct ligations for plasmid construction are confirmed by first transforming *E. coli* strain MM294, or other suitable host, with the ligation mixture. Successful transformants are selected by ampicillin, tetracycline or other antibiotic resistance or using other markers, depending on the mode of plasmid construction, as is understood in the art. Plasmids from the transformants are then prepared according to the method of Clewell, D. B., et al., *Proc. Natl. Acad. Sci.* (U.S.A.) (1969) 62: 1159, optionally following chloramphenicol amplification (Clewell, D. B., *J. Bacteriol.* (1972) 110: 667). The isolated DNA is analyzed by restriction and/or sequenced by the dideoxy method of Sanger, F., et al., *Proc. Natl. Acad. Sci.* (U.S.A.) (1977) 74: 5463 as further described by Messing, et al., *Nucleic Acids Res.* (1981) 9: 309, or by the method of Maxam, et al., *Methods in Enzymology* (1980) 65: 499.

HOST STRAINS EXEMPLIFIED

Host strains used in cloning and expression herein are as follows:

For cloning and sequencing, and for expression of constructions under control of most bacterial promoters, *E. coli* strain MM294 obtained from *E. coli* Genetic Stock Center GCSC #6135, was used as the host. For expression under control of the $P_L N_{RBS}$ promoter, *E. coli* strain K12 MC1000 lambda lysogen, $N_7N_{53}cI857$ $SusP_{80}$, ATCC 39531 may be used. Used herein are *E. coli* DG116, which was deposited with ATCC (ATCC 53606) on Apr. 7, 1987 and *E. coli* KB2, which was deposited with ATCC (ATCC 53075) on Mar. 29, 1985.

For M13 phage recombinants, *E. coli* strains susceptible to phage infection, such as *E. coli* K12 strain DG98, are employed. The DG98 strain has been deposited with ATCC July 13, 1984 and has accession number 39768.

Mammalian expression can be accomplished in COS-7 COS-A2, CV-1, and murine cells, and insect cell-based expression in *Spodoptera frugipeida*).

PURIFICATION

In addition to the purification procedures previously described, the thermostable polymerase of the invention may be purified using hydrophobic interaction chromatography. Hydrophobic interaction chromatography is a separation technique in which substances are separated on the basis of differing strengths of hydrophobic interaction with an uncharged bed material containing hydrophobic groups. Typically, the column is first equilibrated under conditions favorable to hydrophobic binding, e.g., high ionic strength. A descending salt gradient may be used to elute the sample.

According to the invention, the aqueous mixture (containing either native or recombinant polymerase) is loaded onto a column containing a relatively strong hydrophobic gel such as Phenyl Sepharose (manufactured by Pharmacia) or Phenyl TSK (manufactured by Toyo Soda). To promote hydrophobic interaction with a Phenyl Sepharose column, a solvent is used which contains, for example, greater than or equal to 0.2M ammonium sulfate, with 0.2M being preferred. Thus the column and the sample are adjusted to 0.2M ammonium sulfate in 50 mM Tris-1 mM EDTA buffer and the sample applied to the column. The column is washed with the 0.2M ammonium sulfate buffer. The enzyme may then be eluted with solvents which attenuate hydrophobic interactions such as, for example, decreasing salt gradients, ethylene or propylene glycol, or urea. For the recombinant Taq polymerase, a preferred embodiment involves washing the column sequentially with the Tris-EDTA buffer and the Tris-EDTA buffer containing 20% ethylene glycol. The Taq polymerase is subsequently eluted from the column with a 0–4M urea gradient in the Tris-EDTA ethylene glycol buffer.

STABILIZATION OF ENZYME ACTIVITY

For long-term stability, the enzyme herein must be stored in a buffer that contains one or more non-ionic polymeric detergents. Such detergents are generally those that have a molecular weight in the range of approximately 100 to 250,000, preferably about 4,000 to 200,000 daltons and stabilize the enzyme at a pH of from about 3.5 to about 9.5, preferably from about 4 to 8.5. Examples of such detergents include those specified on pages 295–298 of McCutcheon's *Emulsifiers & Detergents*, North American edition (1983), published by the McCutcheon Division of MC Publishing Co., 175 Rock Road, Glen Rock, N.J. (U.S.A.), the entire disclosure of which is incorporated herein by reference. Preferably, the detergents are selected from the group comprising ethoxylated fatty alcohol ethers and lauryl ethers, ethoxylated alkyl phenols, octylphenoxy polyethoxy ethanol compounds, modified oxyethylated and/or oxypropylated straight-chain alcohols, polyethylene glycol monooleate compounds, polysorbate compounds, and phenolic fatty alcohol ethers. More particularly preferred are Tween 20, from ICI Americas Inc., Wilmington, Del., which is a polyoxyethylated (20) sorbitan monolaurate, and Iconol TM NP-40, from BASF Wyandotte Corp. Parsippany, N.J., which is an ethoxylated alkyl phenol (nonyl).

The thermostable enzyme of this invention may be used for any purpose in which such enzyme is necessary or desirable. In a particularly preferred embodiment, the enzyme herein is employed in the amplification protocol set forth below.

AMPLIFICATION PROTOCOL

The amplification protocol using the enzyme of the present invention may be the process for amplifying existing nucleic acid sequences that is disclosed and claimed in U.S. Pat. No. 4,683,202, issued July 28, 1987, the disclosure of which is incorporated herein by reference. Preferably, however, the enzyme herein is used in the amplification process disclosed below.

Specifically, the amplification method involves amplifying at least one specific nucleic acid sequence contained in a nucleic acid or a mixture of nucleic acids, wherein if the nucleic acid is double-stranded, it consists of two separated complementary strands of equal or unequal length, which process comprises:

(a) contacting each nucleic acid strand with four different nucleotide triphosphates and one oligonucleotide primer for each different specific sequence being amplified, wherein each primer is selected to be substantially complementary to different strands of each specific sequence, such that the extension product synthesized from one primer, when it is separated from its complement, can serve as a template for synthesis of the extension product of the other primer, said contacting being at a temperature which promotes hybridization of each primer to its complementary nucleic acid strand;

(b) contacting each nucleic acid strand, at the same time as or after step (a), with a DNA polymerase from *Thermus aquaticus* which enables combination of the nucleotide triphosphates to form primer extension products complementary to each strand of each nucleic acid;

(c) maintaining the mixture from step (b) at an effective temperature for an effective time to promote the activity of the enzyme, and to synthesize, for each different sequence being amplified, an extension product of each primer which is complementary to each nucleic acid strand template, but not so high as to separate each extension product from its complementary strand template;

(d) heating the mixture from step (c) for an effective time and at an effective temperature to separate the primer extension products from the templates on which they were synthesized to produce single-stranded molecules, but not so high as to denature irreversibly the enzyme;

(e) cooling the mixture from step (d) for an effective time and to an effective temperature to promote hybridization of each primer to each of the single-stranded molecules produced in step (d); and

(f) maintaining the mixture from step (e) at an effective temperature for an effective time to promote the activity of the enzyme and to synthesize, for each different sequence being amplified, an extension product of each primer which is complementary to each nucleic acid strand template produced in step (d), but not so high as to separate each extension product from its complementary strand template wherein the effective time and temperatures in steps (e) and (f) may coincide (steps (e) and (f) are carried out simultaneously), or may be separate.

Steps (d)–(f) may be repeated until the desired level of sequence amplification is obtained.

The amplification method is useful not only for producing large amounts of an existing completely specified nucleic acid sequence, but also for producing nucleic acid sequences which are known to exist but are not completely specified. In either case an initial copy

of the sequence to be amplified must be available, although it need not be pure or a discrete molecule.

In general, the amplification process involves a chain reaction for producing, in exponential quantities relative to the number of reaction steps involved, at least one specific nucleic acid sequence given (a) that the ends of the required sequence are known in sufficient detail that oligonucleotides can be synthesized which will hybridize to them, and (b) that a small amount of the sequence is available to initiate the chain reaction. The product of the chain reaction will be a discrete nucleic acid duplex with termini corresponding to the ends of the specific primers employed.

Any nucleic acid sequence, in purified or nonpurified form, can be utilized as the starting nucleic acid(s), provided it contains or is suspected to contain the specific nucleic acid sequence desired. Thus, the process may employ, for example, DNA or RNA, including messenger RNA, which DNA or RNA may be single-stranded or double-stranded. In addition, a DNA-RNA hybrid which contains one strand of each may be utilized. A mixture of any of these nucleic acids may also be employed, or the nucleic acids produced from a previous amplification reaction herein using the same or different primers may be so utilized. The specific nucleic acid sequence to be amplified may be only a fraction of a larger molecule or can be present initially as a discrete molecule, so that the specific sequence constitutes the entire nucleic acid.

It is not necessary that the sequence to be amplified be present initially in a pure form; it may be a minor fraction of a complex mixture, such as a portion of the β-globin gene contained in whole human DNA (as exemplified in Saiki et al., *Science,* 230, 1530–1534 (1985)) or a portion of a nucleic acid sequence due to a particular microorganism which organism might constitute only a very minor fraction of a particular biological sample. The starting nucleic acid sequence may contain more than one desired specific nucleic acid sequence which may be the same or different. Therefore, the amplification process is useful not only for producing large amounts of one specific nucleic acid sequence, but also for amplifying simultaneously more than one different specific nucleic acid sequence located on the same or different nucleic acid molecules.

The nucleic acid(s) may be obtained from any source, for example, from plasmids such as pBR322, from cloned DNA or RNA, or from natural DNA or RNA from any source, including bacteria, yeast, viruses, organelles, and higher organisms such as plants or animals. DNA or RNA may be extracted from blood, tissue material such as chorionic villi, or amniotic cells by a variety of techniques such as that described by Maniatis et al., supra, p. 280–281.

If probes are used which are specific to a sequence being amplified and thereafter detected, the cells may be directly used without extraction of the nucleic acid if they are suspended in hypotonic buffer and heated to about 90°–100° C., until cell lysis and dispersion of intracellular components occur, generally 1 to 15 minutes. After the heating step the amplification reagents may be added directly to the lysed cells.

Any specific nucleic acid sequence can be produced by the amplification process. It is only necessary that a sufficient number of bases at both ends of the sequence be known in sufficient detail so that two oligonucleotide primers can be prepared which will hybridize to different strands of the desired sequence and at relative positions along the sequence such that an extension product synthesized from one primer, when it is separated from its template (complement), can serve as a template for extension of the other primer into a nucleic acid sequence of defined length. The greater the knowledge about the bases at both ends of the sequence, the greater can be the specificity of the primers for the target nucleic acid sequence, and thus the greater the efficiency of the process.

It will be understood that the word "primer" as used hereinafter may refer to more than one primer, particularly in the case where there is some ambiguity in the information regarding the terminal sequence(s) of the fragment to be amplified. For instance, in the case where a nucleic acid sequence is inferred from protein sequence information, a collection of primers containing sequences representing all possible codon variations based on degeneracy of the genetic code will be used for each strand. One primer from this collection will be homologous with the end of the desired sequence to be amplified.

The oligonucleotide primers may be prepared using any suitable method, such as, for example, the phosphotriester and phosphodiester methods described above, or automated embodiments thereof. In one such automated embodiment, diethylphosphoramidites are used as starting materials and may be synthesized as described by Beaucage et al., *Tetrahedron Letters* (1981), 22: 1859–1862. One method for synthesizing oligonucleotides on a modified solid support is described in U.S. Pat. No. 4,458,066. It is also possible to use a primer which has been isolated from a biological source (such as a restriction endonuclease digest).

The specific nucleic acid sequence is produced by using the nucleic acid containing that sequence as a template. The first step involves contacting each nucleic acid strand with four different nucleotide triphosphates and one oligonucleotide primer for each different nucleic acid sequence being amplified or detected. If the nucleic acids to be amplified or detected are DNA, then the nucleotide triphosphates are dATP, dCTP, dGTP and TTP.

The nucleic acid strands are used as a template for the synthesis of additional nucleic acid strands. This synthesis can be performed using any suitable method. Generally it occurs in a buffered aqueous solution, preferably at a pH of 7–9, most preferably about 8. Preferably, a molar excess (for cloned nucleic acid, usually about 1000:1 primer:template, and for genomic nucleic acid, usually about 10^8:1 primer:template) of the two oligonucleotide primers is added to the buffer containing the separated template strands. It is understood, however, that the amount of complementary strand may not be known if the process herein is used for diagnostic applications, so that the amount of primer relative to the amount of complementary strand cannot be determined with certainty. As a practical matter, however, the amount of primer added will generally be in molar excess over the amount of complementary strand (template) when the sequence to be amplified is contained in a mixture of complicated long-chain nucleic acid strands. A large molar excess is preferred to improve the efficiency of the process.

Preferably the concentration of nucleotide triphosphates is 150–200 μM each in the buffer for amplification and $MgCl_2$ is present in the buffer in an amount of 1.5–2 mM to increase the efficiency and specificity of the reaction.

The resulting solution is then treated according to whether the nucleic acids being amplified or detected are double or single-stranded. If the nucleic acids are single-stranded, then no denaturation step need be employed, and the reaction mixture is held at a temperature which promotes hybridization of the primer to its complementary target (template) sequence. Such temperature is generally from about 35° C. to 65° C. or more, preferably about 37°–60° C. for an effective time, generally one-half to five minutes, preferably one-three minutes. Preferably, 45°–58° C. is used for Taq polymerase and >15-mer primers to increase the specificity of primer hybridization. Shorter primers need lower temperatures.

The complement to the original single-stranded nucleic acid may be synthesized by adding one or two oligonucleotide primers thereto. If an appropriate single primer is added, a primer extension product is synthesized in the presence of the primer, the DNA polymerase from *Thermus aquaticus* and the nucleotide triphosphates. The product will be partially complementary to the single-stranded nucleic acid and will hybridize with the nucleic acid strand to form a duplex of strands of unequal length which may then be separated into single strands as described above to produce two single separated complementary strands. Alternatively, two appropriate primers may be added to the single-stranded nucleic acid and the reaction carried out.

If the nucleic acid contains two strands, it is necessary to separate the strands of the nucleic acid before it can be used as the template. This strand separation can be accomplished by any suitable denaturing method including physical, chemical or enzymatic means. One preferred physical method of separating the strands of the nucleic acid involves heating the nucleic acid until it is completely (>99%) denatured. Typical heat denaturation involves temperatures ranging from about 90° to 105° C. for times generally ranging from about 0.5 to 5 minutes. Preferably the effective denaturing temperature is 90°–100° C. for 0.5 to 3 minutes. Strand separation may also be induced by an enzyme from the class of enzymes known as helicases or the enzyme RecA, which has helicase activity and in the presence of riboATP is known to denature DNA. The reaction conditions suitable for separating the strands of nucleic acids with helicases are described by Kuhn Hoffmann-Berling, *CSH-Quantitative Biology,* 43: 63 (1978), and techniques for using RecA are reviewed in C. Radding, *Ann. Rev. Genetics,* 16: 405–37 (1982). The denaturation produces two separated complementary strands of equal or unequal length.

If the double-stranded nucleic acid is denatured by heat, the reaction mixture is allowed to cool to a temperature which promotes hybridization of each primer present to its complementary target (template) sequence. This temperature is usually from about 35° C. to 65° C. or more, depending on reagents, preferably 37°–60° C., maintained for an effective time, generally 0.5 to 5 minutes, and preferably 1–3 minutes. In practical terms, the temperature is simply lowered from about 95° C. to as low as 37° C., preferably to about 45°–58° C. for Taq polymerase, and hybridization occurs at a temperature within this range.

Whether the nucleic acid is single- or double-stranded, the DNA polymerase from *Thermus aquaticus* may be added at the denaturation step or when the temperature is being reduced to or is in the range for promoting hybridization. The reaction mixture is then heated to a temperature at which the activity of the enzyme is promoted or optimized, i.e., a temperature sufficient to increase the activity of the enzyme in facilitating synthesis of the primer extension products from the hybridized primer and template. The temperature must actually be sufficient to synthesize an extension product of each primer which is complementary to each nucleic acid template, but must not be so high as to denature each extension product from its complementary template (i.e., the temperature is generally less than about 80° C.–90° C.).

Depending mainly on the types of enzyme and nucleic acid(s) employed, the typical temperature effective for this synthesis reaction generally ranges from about 40° to 80° C., preferably 50°–75° C. The temperature more preferably ranges from about 65°–75° C. when a DNA polymerase from *Thermus aquaticus* is employed. The period of time required for this synthesis may range from about 0.5 to 40 minutes or more, depending mainly on the temperature, the length of the nucleic acid, the enzyme and the complexity of the nucleic acid mixture, preferably one to three minutes. If the nucleic acid is longer, a longer time period is generally required. The presence of dimethylsulfoxide (DMSO) is not necessary or recommended because DMSO was found to inhibit Taq polymerase enzyme activity.

The newly synthesized strand and its complementary nucleic acid strand form a double-stranded molecule which is used in the succeeding steps of the process. In the next step, the strands of the double-stranded molecule are separated by heat denaturation at a temperature effective to denature the molecule, but not so high that the thermostable enzyme is completely and irreversibly denatured or inactivated. Depending mainly on the type of enzyme and the length of nucleic acid, this temperature generally ranges from about 90° to 105° C., more preferably 90°–100° C., and the time for denaturation typically ranges from 0.5 to four minutes, depending mainly on the temperature and nucleic acid length.

After this time, the temperature is decreased to a level which promotes hybridization of the primer to its complementary single-stranded molecule (template) produced from the previous step. Such temperature is described above.

After this hybridization step, or in lieu of (or concurrently with) the hybridization step, the temperature is adjusted to a temperature that is effective to promote the activity of the thermostable enzyme to enable synthesis of a primer extension product using as template the newly synthesized strand from the previous step. The temperature again must not be so high as to separate (denature) the extension product from its template, as previously described (usually from 40° to 80° C. for 0.5 to 40 minutes, preferably 50° to 70° C. for one-three minutes). Hybridization may occur during this step, so that the previous step of cooling after denaturation is not required. In such a case, using simultaneous steps, the preferred temperature range is 50°–70° C.

The heating and cooling steps of strand separation, hybridization, and extension product synthesis can be repeated as often as needed to produce the desired quantity of the specific nucleic acid sequence, depending on the ultimate use. The only limitation is the amount of the primers, thermostable enzyme and nucleotide triphosphates present. Preferably, the steps are repeated at least twice. For use in detection, the number of cycles will depend, e.g., on the nature of the sample.

For example, fewer cycles will be required if the sample being amplified is pure. If the sample is a complex mixture of nucleic acids, more cycles will be required to amplify the signal sufficiently for its detection. For general amplification and detection, preferably the process is repeated at least 20 times.

When labeled sequence-specific probes are employed as described below, preferably the steps are repeated at least five times. When human genomic DNA is employed with such probes, the process is repeated preferably 15–30 times to amplify the sequence sufficiently that a clearly detectable signal is produced, i.e., so that background noise does not interfere with detection.

As will be described in further detail below, the amount of the specific nucleic acid sequence produced will accumulate in an exponential fashion.

No additional nucleotides, primers, or thermostable enzyme need be added after the initial addition, provided that the enzyme has not become denatured or inactivated irreversibly, in which case it is necessary to replenish the enzyme after each denaturing step. Addition of such materials at each step, however, will not adversely affect the reaction.

When it is desired to produce more than one specific nucleic acid sequence from the first nucleic acid or mixture of nucleic acids, the appropriate number of different oligonucleotide primers are utilized. For example, if two different specific nucleic acid sequences are to be produced, four primers are utilized. Two of the primers are specific for one of the specific nucleic acid sequences and the other two primers are specific for the second specific nucleic acid sequence. In this manner, each of the two different specific sequences can be produced exponentially by the present process.

After the appropriate length of time has passed to produce the desired amount of the specific nucleic acid sequence, the reaction may be halted by inactivating the enzyme in any known manner (e.g., by adding EDTA, phenol, SDS or $CHCl_3$) or by separating the components of the reaction.

The amplification process may be conducted continuously. In one embodiment of an automated process, the reaction mixture may be temperature cycled such that the temperature is programmed to be controlled at a certain level for a certain time.

One such instrument for this purpose is the automated machine for handling the amplification reaction of this invention described in now abandoned Ser. No. 833,368 filed Feb. 25, 1986 entitled "Apparatus And Method For Performing Automated Amplification of Nucleic Acid Sequences And Assays Using Heating And Cooling Steps," the disclosure of which is incorporated herein by reference. Briefly, this instrument utilizes a liquid handling system under computer control to make liquid transfers of enzyme stored at a controlled temperature in a first receptacle into a second receptacle whose temperature is controlled by the computer to conform to a certain incubation profile. The second receptacle stores the nucleic acid sequence(s) to be amplified plus the nucleotide triphosphates and primers. The computer includes a user interface through which a user can enter process parameters that control the characteristics of the various steps in the amplification sequence such as the times and temperatures of incubation, the amount of enzyme to transfer, etc.

A preferred machine that may be employed utilizes temperature cycling without a liquid handling system because the enzyme need not be transferred at every cycle. Such a machine is described more completely in European Patent Application No. 236,069, published Sept. 9, 1987, the disclosure of which is incorporated herein by reference. Briefly, this instrument consists of the following systems:

1. A heat-conducting container for holding a given number of tubes, preferably 500 μl tubes, which contain the reaction mixture of nucleotide triphosphates, primers, nucleic acid sequences, and enzyme.

2. A means to heat, cool, and maintain the heat-conducting container above and below room temperature, which means has an input for receiving a control signal for controlling which of the temperatures at or to which the container is heated, cooled or maintained. (These may be Peltier heat pumps available from Materials Electronics Products Corporation in Trenton, N.J. or a water heat exchanger.)

3. A computer means (e.g., a microprocessor controller), coupled to the input of said means, to generate the signals that control automatically the amplification sequence, the temperature levels, and the temperature ramping and timing.

The amplification protocol is demonstrated diagrammatically below, where double-stranded DNA containing the desired sequence [S] comprised of complementary strands [S+] and [S−] is utilized as the nucleic acid. During the first and each subsequent reaction cycle, extension of each oligonucleotide primer on the original template will produce one new ssDNA molecule product of indefinite length that terminates with only one of the primers. These products, hereafter referred to as "long products," will accumulate in a linear fashion; that is, the amount present after any number of cycles will be proportional to the number of cycles.

The long products thus produced will act as templates for one or the other of the oligonucleotide primers during subsequent cycles and will produce molecules of the desired sequence [S+] or [S−]. These molecules will also function as templates for one or the other of the oligonucleotide primers, producing further [S+] and [S−], and thus a chain reaction can be sustained that will result in the accumulation of [S] at an exponential rate relative to the number of cycles.

By-products formed by oligonucleotide hybridizations other than those intended are not self-catalytic (except in rare instances) and thus accumulate at a linear rate.

The specific sequence to be amplified, [S], can be depicted diagrammatically as:

[S+] 5′ AAAAAAAAAAXXXXXXXXXXCCCCCCCCCC 3′
[S−] 3′ TTTTTTTTTTYYYYYYYYYYGGGGGGGGGG 5′

The appropriate oligonucleotide primers would be:

Primer 1: 3′ GGGGGGGGGG 5′
Primer 2: 5′ AAAAAAAAAA 3′

so that if DNA containing [S]

... zzzzzzzzzzzzzzzzAAAAAAAAAAXXXXXXX-XXXCCCCCCCCCCzzzzzzzzzzzzzzz ...

... zzzzzzzzzzzzzzzzTTTTTTTTTTYYYYYYYYYY-YGGGGGGGGGGzzzzzzzzzzzzzzz ...

is separated into single strands and its single strands are hybridized to Primers 1 and 2, the following extension reactions can be catalyzed by a thermostable polymerase in the presence of the four nucleotide triphosphates:

On denaturation of the two duplexes formed, the products are:

3′ 5′
. . . zzzzzzzzzzzzzzzTTTTTTTTTTGGGGGGGGGG
newly synthesized long product 1

5′ 3′
. . . zzzzzzzzzzzzzzzAAAAAAAAAAXXXXXXXXXXCCCCCCCCCCzzzzzzzzzzzzzzz . . .
original template strand+

3′ 5′
. . . zzzzzzzzzzzzzzzTTTTTTTTTTYYYYYYYYYYGGGGGGGGGGzzzzzzzzzzzzzzz . . .
original template strand−

5′ 3′
AAAAAAAAAAXXXXXXXXXXCCCCCCCCCCzzzzzzzzzzzzzzz . . .
newly synthesized long product 2

If these four strands are allowed to rehybridize with Primers 1 and 2 in the next cycle, the thermostable polymerase will catalyze the following reactions:

3′ 5′
extends ← GGGGGGGGGG Primer 1
. . . zzzzzzzzzzzzzzzAAAAAAAAAAXXXXXXXXXXCCCCCCCCCCzzzzzzzzzzzzzzz . . .
original template strand+

original template strand−
. . . zzzzzzzzzzzzzzz TTTTTTTTTTYYYYYYYYYYGGGGGGGGGGzzzzzzzzzzzzzzz . . .
Primer 2 AAAAAAAAAA → extends
5′ 3′

Primer 2 5′ AAAAAAAAAA → extends to here
3′ . . . zzzzzzzzzzzzzzzTTTTTTTTTTYYYYYYYYYYGGGGGGGGGG 5′
newly synthesized long product 1

extends ← GGGGGGGGGG 5′ Primer 1
5′ . . . zzzzzzzzzzzzzzzAAAAAAAAAAXXXXXXXXXXCCCCCCCCCCzzzzzzzzzzzzzzz . . . 3′
original template strand+

Primer 2 5′ AAAAAAAAAA → extends
3′ . . . zzzzzzzzzzzzzzzTTTTTTTTTTYYYYYYYYYYGGGGGGGGGGzzzzzzzzzz . . . 5′
original template strand−

extends to here ← GGGGGGGGGG 5′ Primer 1
5′ AAAAAAAAAAXXXXXXXXXXCCCCCCCCCCzzzzzzzzzzzzzzz . . . 3′
newly synthesized long product 2

If the strands of the above four duplexes are separated, the following strands are found:

5′ AAAAAAAAAAXXXXXXXXXXCCCCCCCCCC 3′
newly synthesized [S+]

3′ . . . zzzzzzzzzzzzzzzzzzTTTTTTTTTTYYYYYYYYYYGGGGGGGGGG 5′
first cycle synthesized long product 1

3′ . . . zzzzzzzzzzzzzzzzzzTTTTTTTTTTYYYYYYYYYYGGGGGGGGGG 5′
newly synthesized long product 1

5′ . . . zzzzzzzzzzzzzzzzzzAAAAAAAAAAXXXXXXXXXXCCCCCCCCCCzzzzzzzzz . . . 3′
original template strand+

5′ AAAAAAAAAAXXXXXXXXXXCCCCCCCCCCzzzzzzzzzzzzzzz . . . 3′
newly synthesized long product 2

3′ . . . zzzzzzzzzzzzzzzTTTTTTTTTTYYYYYYYYYYGGGGGGGGGGzzzzzzzzzzzzzzz . . . 5′
original template strand−

Case 3:11-cv-00056-IEG -WVG Document 1-2 Filed 01/10/11 Page 22 of 90

-continued

3′ TTTTTTTTTTYYYYYYYYYYGGGGGGGGGG 5′
newly synthesized [S⁻]

5′ AAAAAAAAAAXXXXXXXXXXCCCCCCCCCCzzzzzzzzzzzzzz . . . 3′
first cycle synthesized long product 2

It is seen that each strand which terminates with the oligonucleotide sequence of one primer and the complementary sequence of the other is the specific nucleic acid sequence [S] that is desired to be produced.

The amount of original nucleic acid remains constant in the entire process, because it is not replicated. The amount of the long products increases linearly because they are produced only from the original nucleic acid. The amount of the specific sequence increases exponentially. Thus, the specific sequence will become the predominant species. This is illustrated in the following table, which indicates the relative amounts of the species theoretically present after n cycles, assuming 100% efficiency at each cycle:

	Number of Double Strands After 0 to n Cycles		
Cycle Number	Template	Long Products	Specific Sequence [S]
0	1	—	—
1	1	1	0
2	1	2	1
3	1	3	4
5	1	5	26
10	1	10	1013
15	1	15	32,752
20	1	20	1,048,555
n	1	n	$(2^n\text{-}n\text{-}1)$

When a single-stranded nucleic acid is utilized as the template, only one long product is formed per cycle.

A sequence within a given sequence can be amplified after a given number of amplifications to obtain greater specificity of the reaction by adding after at least one cycle of amplification a set of primers that are complementary to internal sequences (that are not on the ends) of the sequence to be amplified. Such primers may be added at any stage and will provide a shorter amplified fragment. Alternatively, a longer fragment can be prepared by using primers with non-complementary ends but having some overlap with the primers previously utilized in the amplification.

The amplification method may be utilized to clone a particular nucleic acid sequence for insertion into a suitable expression vector. The vector may be used to transform an appropriate host organism to produce the gene product of the sequence by standard methods of recombinant DNA technology. Such cloning may involve direct ligation into a vector using blunt-end ligation, or use of restriction enzymes to cleave at sites contained within the primers.

In addition, the amplification process can be used for in vitro mutagenesis. The oligodeoxyribonucleotide primers need not be exactly complementary to the DNA sequence that is being amplified. It is only necessary that they be able to hybridize to the sequence sufficiently well to be extended by the thermostable enzyme. The product of an amplification reaction wherein the primers employed are not exactly complementary to the original template will contain the sequence of the primer rather than the template, thereby introducing an in vitro mutation. In further cycles this mutation will be amplified with an undiminished efficiency because no further mispaired priming is required. The mutant thus produced may be inserted into an appropriate vector by standard molecular biological techniques and might confer mutant properties on this vector such as the potential for production of an altered protein.

The process of making an altered DNA sequence as described above could be repeated on the altered DNA using different primers to induce further sequence changes. In this way, a series of mutated sequences could gradually be produced wherein each new addition to the series could differ from the last in a minor way, but from the original DNA source sequence in an increasingly major way. In this manner, changes could be made ultimately which were not feasible in a single step due to the inability of a very seriously mismatched primer to function.

In addition, the primer can contain as part of its sequence a non-complementary sequence, provided that a sufficient amount of the primer contains a sequence that is complementary to the strand to be amplified. For example, a nucleotide sequence that is not complementary to the template sequence (such as, e.g., a promoter, linker, coding sequence, etc.) may be attached at the 5′ end of one or both of the primers, and thereby appended to the product of the amplification process. After the extension primer is added, sufficient cycles are run to achieve the desired amount of new template containing the non-complementary nucleotide insert. This allows production of large quantities of the combined fragments in a relatively short period of time (e.g., two hours or less) using a simple technique.

The amplification method may also be used to enable detection and/or characterization of specific nucleic acid sequences associated with infectious diseases, genetic disorders or cellular disorders such as cancer, e.g., oncogenes. Amplification is useful when the amount of nucleic acid available for analysis is very small, as, for example, in the prenatal diagnosis of sickle cell anemia using DNA obtained from fetal cells. Amplification is particularly useful if such an analysis is to be done on a small sample using non-radioactive detection techniques which may be inherently insensitive, or where radioactive techniques are being employed, but where rapid detection is desirable.

For the purposes of this discussion, genetic diseases may include specific deletions and/or mutations in genomic DNA from any organism, such as, e.g., sickle cell anemia, cystic fibrosis, α-thalassemia, β-thalassemia, and the like. Sickle cell anemia can be readily detected via oligomer restriction analysis as described by EP Patent Publication 164,054 published Dec. 11, 1985, or via a RFLP-like analysis following amplification of the appropriate DNA sequence by the amplification method. α-Thalassemia can be detected by the absence of a sequence, and β-thalassemia can be detected by the presence of a polymorphic restriction site closely linked to a mutation that causes the disease.

All of these genetic diseases may be detected by amplifying the appropriate sequence and analyzing it by Southern blots without using radioactive probes. In

such a process, for example, a small sample of DNA from, e.g., amniotic fluid containing a very low level of the desired sequence is amplified, cut with a restriction enzyme, and analyzed via a Southern blotting technique. The use of non-radioactive probes is facilitated by the high level of the amplified signal.

In another embodiment, a small sample of DNA may be amplified to a convenient level and then a further cycle of extension reactions performed wherein nucleotide derivatives which are readily detectable (such as ^{32}P-labeled or biotin-labeled nucleotide triphosphates) are incorporated directly into the final DNA product, which may be analyzed by restriction and electrophoretic separation or any other appropriate method.

In a further embodiment, the nucleic acid may be exposed to a particular restriction endonuclease prior to amplification. Since a sequence which has been cut cannot be amplified, the appearance of an amplified fragment, despite prior restriction of the DNA sample, implies the absence of a site for the endonuclease within the amplified sequence. The presence or absence of an amplified sequence can be detected by an appropriate method.

A practical application of the amplification technique, that is, in facilitating the detection of sickle cell anemia via the oligomer restriction technique [described in EP 164,054, supra, and by Saiki et al., *Bio/-Technology*, Vol. 3, pp. 1008–1012 (1985)] is described in detail in the Saiki et al. *Science* article cited above. In that *Science* article, a specific amplification protocol is exemplified using a β-globin gene segment.

The amplification method herein may also be used to detect directly single-nucleotide variations in nucleic acid sequence (such as genomic DNA) using sequence-specific oligonucleotides, as described more fully in European Patent Publication 237,362, published Sept. 16, 1987, the disclosure of which is incorporated herein by reference.

Briefly, in this process, the amplified sample is spotted directly on a series of membranes, and each membrane is hybridized with a different labeled sequence-specific oligonucleotide probe. After hydridization the sample is washed and the label is detected. This technique is especially useful in detecting DNA polymorphisms.

Various infectious diseases can be diagnosed by the presence in clinical samples of specific DNA sequences characteristic of the causative microorganism. These include bacteria, such as Salmonella, Chlamydia, Neisseria; viruses, such as the hepatitis viruses, and parasites, such as the Plasmodium responsible for malaria. U.S. Patent Reexamination Certificate B1 4,358,535 issued to Falkow et al. on May 13, 1986 describes the use of specific DNA hybridization probes for the diagnosis of infectious diseases. A relatively small number of pathogenic organisms may be present in a clinical sample from an infected patient and the DNA extracted from these may constitute only a very small fraction of the total DNA in the sample. Specific amplification of suspected pathogen-specific sequences prior to immobilization and detection by hybridization of the DNA samples could greatly improve the sensitivity and specificity of traditional procedures.

Routine clinical use of DNA probes for the diagnosis of infectious diseases would be simplified considerably if non-radioactively labeled probes could be employed as described in EP 63,879 to Ward. In this procedure biotin-containing DNA probes are detected by chromogenic enzymes linked to avidin or biotin-specific antibodies. This type of detection is convenient, but relatively insensitive. The combination of specific DNA amplification by the present method and the use of stably labeled probes could provide the convenience and sensitivity required to make the Falkow et al. and Ward procedures useful in a routine clinical setting.

A specific use of the amplification technology for detecting or monitoring for the AIDS virus is described in European Patent Publication 229,701, published July 22, 1987, the disclosure of which is incorporated herein by reference. Briefly, the amplification and detection process is used with primers and probes which are designed to amplify and detect, respectively, nucleic acid sequences that are substantially conserved among the nucleic acids in AIDS viruses and specific to the nucleic acids in AIDS viruses. Thus, the sequence to be detected must be sufficiently complementary to the nucleic acids in AIDS viruses to initiate polymerization preferably at room temperature in the presence of the enzyme and nucleotide triphosphates.

A preferred amplification process described in U.S. Ser. No. 07/076,394, filed July 22, 1987, assigned to the same assignee, and incorporated herein by reference, uses labeled primers. The label on the amplified product may be used to "capture" or immobilize the product for subsequent detection (e.g., biotinylated amplification primers yield labeled products that can be "captured" by their interaction with avidin or strepavidin). As demonstrated in the aforementioned amplification protocols, the extension product of one labeled primer when hybridized to the other becomes a template for the production of the desired specific nucleic acid sequence, and vice versa, and the process is repeated as often as necessary to produce the desired amount of the sequence. Examples of specific preferred reagents that can be employed as the label are provided in U.S. Pat. No. 4,582,789, the disclosure of which is incorporated herein by reference.

The amplification process can also be utilized to produce sufficient quantities of DNA from a single copy human gene such that detection by a simple non-specific DNA stain such as ethidium bromide can be employed to diagnose DNA directly.

In addition to detecting infectious diseases and pathological abnormalities in the genome of organisms, the amplification process can also be used to detect DNA polymorphisms which may not be associated with any pathological state.

In summary, the amplification process is seen to provide a process for amplifying one or more specific nucleic acid sequences using a chain reaction and a thermostable enzyme, in which reaction primer extension products are produced which can subsequently act as templates for further primer extension reactions. The process is especially useful in detecting nucleic acid sequences which are initially present in only very small amounts.

The following examples are offered by way of illustration only and are by no means intended to limit the scope of the claimed invention. In these examples, all percentages are by weight if for solids and by volume if for liquids, unless otherwise noted, and all temperatures are given in degrees Celsius.

EXAMPLE I

A. Synthesis of the Primers

The following two oligonucleotide primers were prepared by the method described below:

5'-ACACAACTGTGTTCACTAGC-3'(PC03)

5'-CAACTTCATCCACGTTCACC-3'(PC04)

These primers, both 20-mers, anneal to opposite strands of the genomic DNA with their 5' ends separated by a distance of 110 base pairs.

1. Automated Synthesis Procedures: The diethylphosphoramidites, synthesized according to Beaucage and Caruthers (*Tetrahedron Letters* (1981) 22:1859–1862) were sequentially condensed to a nucleoside derivatized controlled pore glass support using a Biosearch SAM-1. The procedure included detritylation with trichloroacetic acid in dichloromethane, condensation using benzotriazole as activating proton donor, and capping with acetic anhydride and dimethylaminopyridine in tetrahydrofuran and pyridine. Cycle time was approximately 30 minutes. Yields at each step were essentially quantitative and were determined by collection and spectroscopic examination of the dimethoxytrityl alcohol released during detritylation.

2. Oligodeoxyribonucleotide Deprotection and Purification Procedures: The solid support was removed from the column and exposed to 1 ml concentrated ammonium hydroxide at room temperature for four hours in a closed tube. The support was then removed by filtration and the solution containing the partially protected oligodeoxynucleotide was brought to 55° C. for five hours. Ammonia was removed and the residue was applied to a preparative polyacrylamide gel. Electrophoresis was carried out at 30 volts/cm for 90 minutes after which the band containing the product was identified by UV shadowing of a fluorescent plate. The band was excised and eluted with 1 ml distilled water overnight at 4° C. This solution was applied to an Altech RP18 column and eluted with a 7–13% gradient of acetonitrile in 1% ammonium acetate buffer at pH 6.0. The elution was monitored by UV absorbance at 260 nm and the appropriate fraction collected, quantitated by UV absorbance in a fixed volume and evaporated to dryness at room temperature in a vacuum centrifuge.

3. Characterization of Oligodeoxyribonucleotides: Test aliquots of the purified oligonucleotides were ^{32}P labeled with polynucleotide kinase and γ-^{32}P-ATP. The labeled compounds were examined by autoradiography of 14–20% polyacrylamide gels after electrophoresis for 45 minutes at 50 volts/cm. This procedure verifies the molecular weight. Base composition was determined by digestion of the oligodeoxyribonucleotide to nucleosides by use of venom diesterase and bacterial alkaline phosphatase and subsequent separation and quantitation of the derived nucleosides using a reverse phase HPLC column and a 10% acetonitrile, 1% ammonium acetate mobile phase.

B. Isolation of Human Genomic DNA from Cell Line

High molecular weight genomic DNA was isolated from a T cell line, Molt 4, homozygous for normal β-globin available from the Human Genetic Mutant Cell Depository, Camden, N.J. as GM2219C using essentially the method of Maniatis et al., supra, p. 280–281.

C. Purification of a Polymerase From *Thermus aquaticus*

Thermus aquaticus strain YT1, available without restriction from the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Md., as ATCC No. 25,104 was grown in flasks in the following medium:

Sodium Citrate	1 mM
Potassium Phosphate, pH 7.9	5 mM
Ammonium Chloride	10 mM
Magnesium Sulfate	0.2 mM
Calcium Chloride	0.1 mM
Sodium Chloride	1 g/l
Yeast Extract	1 g/l
Tryptone	1 g/l
Glucose	2 g/l
Ferrous Sulfate	0.01 mM

(The pH was adjusted to 8.0 prior to autoclaving.)

A 10-liter fermentor was inoculated from a seed flask cultured overnight in the above medium at 70° C. A total of 600 ml from the seed flask was used to inoculate 10 liters of the same medium. The pH was controlled at 8.0 with ammonium hydroxide with the dissolved oxygen at 40%, with the temperature at 70° C., and with the stirring rate at 400 rpm.

After growth of the cells, they were purified using the protocol (with slight modification) of Kaledin et al., supra, through the first five stages and using a different protocol for the sixth stage. All six steps were conducted at 4° C. The rate of fractionation on columns was 0.5 columns/hour and the volumes of gradients during elution were 10 column volumes. An alternative and preferred purification protocol is provided in Example XIII below.

Briefly, the above culture of the *T. aquaticus* cells was harvested by centrifugation after nine hours of cultivation, in late log phase, at a cell density of 1.4 g dry weight/l. Twenty grams of cells were resuspended in 80 ml of a buffer consisting of 50 mM Tris.HCl pH 7.5, 0.1 mM EDTA. Cells were lysed and the lysate was centrifuged for two hours at 35,000 rpm in a Beckman TI 45 rotor at 4° C. The supernatant was collected (fraction A) and the protein fraction precipitating between 45 and 75% saturation of ammonium sulfate was collected, dissolved in a buffer consisting of 0.2M potassium phosphate buffer, pH 6.5, 10 mM 2-mercaptoethanol, and 5% glycerine, and finally dialyzed against the same buffer to yield fraction B.

Fraction B was applied to a 2.2×30-cm column of DEAE-cellulose, equilibrated with the above described buffer. The column was then washed with the same buffer and the fractions containing protein (determined by absorbance at 280 nm) were collected. The combined protein fraction was dialyzed against a second buffer, containing 0.01M potassium phosphate buffer, pH 7.5, 10 mM 2-mercaptoethanol, and 5% glycerine, to yield fraction C.

Fraction C was applied to a 2.6×21-cm column of hydroxyapatite, equilibrated with a second buffer. The column was then washed and the enzyme was eluted with a linear gradient of 0.01–0.5M potassium phosphate buffer, pH 7.5, containing 10 mM 2-mercaptoethanol and 5% glycerine. Fractions containing DNA polymerase activity (90–180 mM potassium phosphate)

were combined, concentrated four-fold using an Amicon stirred cell and YM10 membrane, and dialyzed against the second buffer to yield fraction D.

Fraction D was applied to a 1.6×28-cm column of DEAE-cellulose, equilibrated with the second buffer. The column was washed and the polymerase was eluted with a linear gradient of 0.01–0.5M potassium phosphate in the second buffer. The fractions were assayed for contaminating endonuclease(s) and exonuclease(s) by electrophoretically detecting the change in molecular weight of phage λ DNA or supercoiled plasmid DNA after incubation with an excess of DNA polymerase (for endonuclease) and after treatment with a restriction enzyme that cleaves the DNA into several fragments (for exonuclease). Only those DNA polymerase fractions (65–95 mM potassium phosphate) having minimal nuclease contamination were pooled. To the pool was added autoclaved gelatin in an amount of 250 μg/ml, and dialysis was conducted against the second buffer to yield Fraction E.

Fraction E was applied to a phosphocellulose column and eluted with a 100 ml gradient (0.01–0.4M KCl gradient in 20 mM potassium phosphate buffer pH 7.5). The fractions were assayed for contaminating endo/exonuclease(s) as described above as well as for polymerase activity (by the method of Kaledin et al.) and then pooled. The pooled fractions were dialyzed against the second buffer, then concentrated by dialysis against 50% glycerine and the second buffer.

The molecular weight of the polymerase was determined by SDS-PAGE analysis. Marker proteins (Bio-Rad low molecular weight standards) were phosphorylase B (92,500), bovine serum albumin (66,200), ovalbumin (45,000), carbonic anhydrase (31,000), soybean trypsin inhibitor (21,500), and lysozyme (14,400).

Preliminary data suggest that the polymerase has a molecular weight of about 86,000–90,000 daltons, not 62,000–63,000 daltons reported in the literature (e.g., by Kaledin et al.).

The polymerase was incubated in 50 μl of a mixture containing either 25 mM Tris-HCl pH 6.4 or pH 8.0, and 0.1M KCl, 10 mM $MgCl_2$, 1 mM 2-mercaptoethanol, 10 nmoles each of dGTP, dATP, and TTP, and 0.5 μCi (^3H) dCTP, 8 μg "activated" calf thymus DNA, and 0.5–5 units of the polymerase. "Activated" DNA is a native preparation of DNA after partial hydrolysis with DNase I until 5% of the DNA was transferred to the acid-soluble fraction. The reaction was conducted at 70° C. for 30 minutes, and stopped by adding 50 μl of a saturated aqueous solution of sodium pyrophosphate containing 0.125M EDTA-Na_2. Samples were processed and activity was determined as described by Kaledin et al., supra.

The results showed that at pH 6.4 the polymerase was more than one-half as active as at pH 8.0. In contrast, Kaledin et al. found that at pH about 7.0, the enzyme therein had 8% of the activity at pH 8.3. Therefore, the pH profile for the thermostable enzyme herein is broader than that for the Kaledin et al. enzyme.

Finally, when only one or more nucleotide triphosphates were eliminated from a DNA polymerase assay reaction mixture, very little, if any, activity was observed using the enzyme herein, and the activity was consistent with the expected value, and with an enzyme exhibiting high fidelity. In contrast, the activity observed using the Kaledin et al. (supra) enzyme is not consistent with the expected value, and suggests misincorporation of nucleotide triphosphate(s).

D. Amplification Reaction

One microgram of the genomic DNA described above was diluted in an initial 100 μl aqueous reaction volume containing 25 mM Tris.HCl buffer (pH 8.0), 50 mM KCl, 10 mM $MgCl_2$, 5 mM dithiothreitol, 200 μg/ml gelatin, 1 μM of primer PC03, 1 μM of primer PC04, 1.5 mM dATP, 1.5 mM dCTP, 1.5 mM dGTP and 1.5 mM TTP. The sample was heated for 10 minutes at 98° C. to denature the genomic DNA, then cooled to room temperature. Four microliters of the polymerase from *Thermus aquaticus* was added to the reaction mixture and overlaid with a 100 μl mineral oil cap. The sample was then placed in the aluminum heating block of the liquid handling and heating instrument described in now abandoned Ser. No. 833,368 filed Feb. 25, 1986, the disclosure of which is incorporated herein by reference.

The DNA sample underwent 20 cycles of amplification in the machine, repeating the following program cycle:

1) heating from 37° C. to 98° C. in heating block over a period of 2.5 minutes; and

2) cooling from 98° C. to 37° C. over a period of three minutes to allow the primers and DNA to anneal.

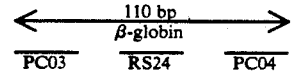
After the last cycle, the sample was incubated for an additional 10 minutes at 55° C. to complete the final extension reaction.

E. Synthesis and Phosphorylation of Oligodeoxyribonucleotide Probes

A labeled DNA probe, designated RS24, of the following sequence was prepared:

5'-*CCCACAGGGCAGTAACG-
GCAGACTTCTCCTCAGGAGTCAG-3' (RS24)

where * indicates the label. This probe is 40 bases long, spans the fourth through seventeenth condons of the gene, and is complementary to the normal β-globin allele (β^A). The schematic diagram of primers and probes is given below:



This probe was synthesized according to the procedures described in Section I of Example I. The probe was labeled by contacting 20 pmole thereof with 4 units of T4 polynucleotide kinase (New England Biolabs) and about 40 pmole γ-^{32}P-ATP (New England Nuclear, about 7000 Ci/mmole) in a 40 μl reaction volume containing 70 mM Tris buffer (pH 7.6), 10 mM $MgCl_2$, 1.5 mM spermine, and 10 mM dithiothreitol for 60 minutes at 37° C. The total volume was then adjusted to 100 μl with 25 mM EDTA and the probe purified according to the procedure of Maniatis et al., *Molecular Cloning* (1982), 466–467 over a 1 ml Bio Gel P-4 (BioRad) spin dialysis column equilibrated with Tris-EDTA (TE) buffer (10 mM Tris buffer, 0.1 mM EDTA, pH 8.0). TCA precipitation of the reaction product indicated that for RS24 the specific activity was 4.3 μCi/pmole and the final concentration was 0.118 pmole/μl.

F. Dot Blot Hybridizations

Four microliters of the amplified sample from Section IV and 5.6 μl of appropriate dilutions of β-globin plas-

mid DNA calculated to represent amplification efficiencies of 70, 75, 80, 85, 90, 95 and 100% were diluted with 200 μl 0.4N NaOH, 25 mM EDTA and spotted onto a Genatran 45 (Plasco) nylon filter by first wetting the filter with water, placing it in a Bio-Dot (Bio-Rad, Richmond, Calif.) apparatus for preparing dot blots which holds the filters in place, applying the samples, and rinsing each well with 0.1 ml of 20×SSPE (3.6M NaCl, 200 mM NaH_2PO_4, 20 mM EDTA), as disclosed by Reed and Mann, *Nucleic Acids Research*, 13, 7202–7221 (1985). The filters were then removed, rinsed in 20×SSPE, and baked for 30 minutes at 80° C. in a vacuum oven.

After baking, each filter was then contacted with 16 ml of a hybridization solution consisting of 3×SSPE, 5×Denhardt's solution (1×=0.02% polyvinylpyrrolidone, 0.02% Ficoll, 0.02% bovine serum albumin, 0.2 mM Tris, 0.2 mM EDTA, pH 8.0), 0.5% SDS and 30% formamide, and incubated for two hours at 42° C. Then 2 pmole of probe RS24 was added to the hybridization solution and the filter was incubated for two minutes at 42° C.

Finally, each hybridized filter was washed twice with 100 ml of 2×SSPE and 0.1% SDS for 10 minutes at room temperature. Then the filters were treated once with 100 ml of 2×SSPE, 0.1% SDS at 60° C. for 10 minutes.

Each filter was then autoradiographed, with the signal readily apparent after two hours.

G. Discussion of Autoradiogram

The autoradiogram of the dot blots was analyzed after two hours and compared in intensity to standard serial dilution β-globin reconstructions prepared with HaeIII/MaeI-digested pBR:β^A, where β^A is the wild-type allele, as described in Saiki et al., *Science*, supra. Analysis of the reaction product indicated that the overall amplification efficiency was about 95%, corresponding to a 630,000-fold increase in the β-globin target sequence.

EXAMPLE II

A. Amplification Reaction

Two 1 μg samples of genomic DNA extracted from the Molt 4 cell line as described in Example I were each diluted in a 100 μl reaction volume containing 50 mM KCl, 25 mM Tris.HCl buffer pH 8.0, 10 mM $MgCl_2$, 1 μM of primer PC03, 1 μM of primer PC04, 200 μg/ml gelatin, 10% dimethylsulfoxide (by volume), and 1.5 mM each of dATP, dCTP, dGTP and TTP. After this mixture was heated for 10 minutes at 98° C. to denature the genomic DNA, the samples were cooled to room temperature and 4 μl of the polymerase from *Thermus aquaticus* described in Example I was added to each sample. The samples were overlaid with mineral oil to prevent condensation and evaporative loss.

One of the samples was placed in the heating block of the machine described in Example I and subjected to 25 cycles of amplification, repeating the following program cycle:

(1) heating from 37° to 93° C. over a period of 2.5 minutes;

(2) cooling from 93° C. to 37° C. over a period of three minutes to allow the primers and DNA to anneal; and

(3) maintaining at 37° C. for two minutes.

After the last cycle the sample was incubated for an additional 10 minutes at 60° C. to complete the final extension reaction.

The second sample was placed in the heat-conducting container of the machine, described in more detail in EP 236,069, supra. The heat-conducting container is attached to Peltier heat pumps which adjust the temperature upwards or downwards and a microprocessor controller to control automatically the amplification sequence, the temperature levels, the temperature ramping and the timing of the temperature.

The second sample was subjected to 25 cycles of amplification, repeating the following program cycle:

(1) heating from 37° to 95° C. over a period of three minutes;

(2) maintaining at 95° C. for 0.5 minutes to allow denaturation to occur;

(3) cooling from 95° to 37° C. over a period of one minute; and

(4) maintaining at 37° C. for one minute.

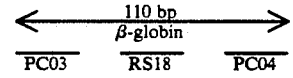
B. Analysis

Two tests were done for analysis, a dot blot and an agarose gel analysis.

For the dot blot analysis, a labeled DNA probe, designated RS18, of the following sequence was prepared.

5'-*CTCCTGAGGAGAAGTCTGC-3'(RS18)

where * indicates the label. This probe is 19 bases long, spans the fourth through seventeenth codons of the gene, and is complementary to the normal β-globin allel (β^A). The schematic diagram of primers and probes is given below:



This probe was synthesized according to the procedures described in Section I of Example I. The probe was labeled by contacting 10 pmole thereof with 4 units of T4 polynucleotide kinase (New England Biolabs) and about 40 pmole γ^{-32}P-ATP (New England Nuclear, about 7000 Ci/mmole) in a 40 μl reaction volume containing 70 mM Tris.HCl buffer (pH 7.6), 10 mM $MgCl_2$, 1.5 mM spermine and 10 mM dithiothreitol for 60 minutes at 37° C. The total volume was then adjusted to 100 μl with 25 mM EDTA and purified according to the procedure of Maniatis et al., supra, p. 466–467 over a 1 ml Bio Gel P-4 (BioRad) spin dialysis column equilibrated with Tris-EDTA (TE) buffer (10 mM Tris.HCl buffer, 0.1 mM EDTA, pH 8.0). TCA precipitation of the reaction product indicated that for RS18 the specific activity was 4.6 μCi/pmole and the final concentration was 0.114 pmole/μl.

Five microliters of the amplified sample from Section I and of a sample amplified as described above except using the Klenow fragment of *E. coli* DNA Polymerase I instead of the thermostable enzyme were diluted with 195 μl 0.4N NaOH, 25 mM EDTA and spotted onto two replicated Genatran 45 (Plasco) nylon filters by first wetting the filters with water, placing them in a Bio-Dot (Bio-Rad, Richmond, Calif.) apparatus for preparing dot blots which holds the filters in place, applying the samples, and rinsing each well with 0.4 ml of 20×SSPE (3.6M NaCl, 200 mM NaH_2PO_4, 20 mM

EDTA), as disclosed by Reed and Mann, supra. The filters were then removed, rinsed in 20×SSPE, and baked for 30 minutes at 80° C. in a vacuum oven.

After baking, each filter was then contacted with 6 ml of a hybridization solution consisting of 5×SSPE, 5×Denhardt's solution (1×=0.02% polyvinylpyrrolidone, 0.02% Ficoll, 0.02% bovine serum albumin, 0.2 mM Tris, 0.2 mM EDTA, pH 8.0) and 0.5% SDS, and incubated for 60 minutes at 55° C. Then 5 μl of probe RS18 was added to the hybridization solution and the filter was incubated for 60 minutes at 55° C.

Finally, each hybridized filter was washed twice with 100 ml of 2×SSPE and 0.1% SDS for 10 minutes at room temperature. Then the filters were treated twice more with 100 ml of 5×SSPE, 0.1% SDS at 60° C. for 1) one minute and 2) three minutes, respectively.

Each filter was then autoradiographed, with the signal readily apparent after 90 minutes.

In the agarose gel analysis, 5 μl of each amplification reaction was loaded onto 4% NuSieve/0.5% agarose gel in 1×TBE buffer (0.089M Tris, 0.089M boric acid, and 2 mM EDTA) and electrophoresed for 60 minutes at 100V. After staining with ethidium bromide, DNA was visualized by UV fluorescence.

The results show that the machines used in Example I and this example were equally effective in amplifying the DNA, showing discrete high-intensity 110-base pair bands of similar intensity, corresponding to the desired sequence, as well as a few other discrete bands of much lower intensity. In contrast, the amplification method as described in Example I of now abandoned Ser. No. 839,331 filed Mar. 13, 1986, supra, which involves reagent transfer after each cycle using the Klenow fragment of *E. coli* Polymerase I, gave a DNA smear resulting from the non-specific amplification of many unrelated DNA sequences.

It is expected that similar improvements in amplification and detection would be achieved in evaluating HLA-DQ, DR and DP regions.

If in the above experiments the amplification reaction buffer contains 2 mM $MgCl_2$ instead of 10 mM $MgCl_2$ and 150–200 μM of each nucleotide rather than 1.5 mM of each, and if the lower temperature of 37° C. is raised to 45°–58° C. during amplification, better specificity and efficiency of amplification occur. Also, DMSO was found not necessary or preferred for amplification.

EXAMPLE III

Amplification and Cloning

For amplification of a 119-base pair fragment on the human β-globin gene, a total of 1 microgram each of human genomic DNA isolated from the Molt 4 cell line or from the GM2064 cell line (representing a homozygous deletion of the β- and δ-hemoglobin region and available from the Human Genetic Mutant Cell Depository, Camden, N.J.) as described above was amplified in a 100 μl reaction volume containing 50 mM KCl, 25 mM Tris.HCl pH 8, 10 mM $MgCl_2$, 200 μg/ml gelatin, 5 mM 2-mercaptoethanol, 1.5 mM each of dATP, dCTP, TTP, and dGTP, and 1 μM of each of the following primers:

5'-CTTCTGcagCAACTGTGTTCACTAGC-3' (GH18)

5'-CACaAgCTTCATCCACGTTCACC-3' (GH19)

where lower case letters denote mismatches from wild-type sequence to create restriction enzyme sites. GH18 is a 26-base oligonucleotide complementary to the negative strand and contains an internal PstI site. GH19 is a 23-base oligonucleotide complementary to the plus strand and contains an internal HindIII recognition sequence. These primers were selected by first screening the regions of the gene for homology to the PstI and HindIII restriction sites. The primers were then prepared as described in Example I.

The above reaction mixtures were heated for 10 minutes at 95° C. and then cooled to room temperature. A total of 4 μl of the polymerase described in Example I was added to each reaction mixture, and then each mixture was overlayed with mineral oil. The reaction mixtures were subjected to 30 cycles of amplification with the following program:

2.5 min. ramp, 37° to 98° C.
3 min. ramp, 98° to 37° C.
2 min. soak, 37° C.

After the last cycle, the reaction mixtures were incubated for 20 minutes at 65° C. to complete the final extension. The mineral oil was extracted with chloroform and the mixtures were stored at −20° C.

A total of 10 μl of the amplified product was digested with 0.5 μg M13mp10 cloning vector, which is publicly available from Boehringer-Mannheim, in a 50 μl volume containing 50 mM NaCl, 10 mM Tris.HCl, pH 7.8, 10 mM $MgCl_2$, 20 units PstI and 26 units HindIII for 90 minutes at 37° C. The reaction was stopped by freezing at −20° C. The volume was adjusted to 110 μl with TE buffer and loaded (100 μl) onto a 1 ml BioGel P-4 spin dialysis column. One 0.1 ml fraction was collected and ethanol precipitated.

(At this point it was discovered that there was β-globin amplification product in the GM2064 sample. Subsequent experiments traced the source of contamination to the primers, either GH18 or GH19. Because no other source of primers was available, the experiment was continued with the understanding that some cloned sequences would be derived from the contaminating DNA in the primers.)

The ethanol pellet was resuspended in 15 μl water, then adjusted to 20 μl volume containing 50 mM Tris.HCl, pH 7.8, 10 mM $MgCl_2$, 0.5 mM ATP, 10 mM dithiothreitol, and 400 units ligase. This mixture was incubated for three hours at 16° C.

Ten microliters of ligation reaction mixture containing Molt 4 DNA was transformed into *E. coli* strain JM103 competent cells, which are publicly available from BRL in Bethesda, MD. The procedure followed for preparing the transformed strain is described in Messing, J. (1981) *Third Cleveland Symposium on Macromolecules:Recombinant DNA*, ed. A. Walton, Elsevier, Amsterdam, 143–163. A total of 651 colorless plaques (and O blue plaques) were obtained. Of these, 119 had a (+)-strand insert (18%) and 19 had a (−)-strand insert (3%). This is an increase of almost 20-fold over the percentage of β-globin positive plaques among the primer-positive plaques from the amplification technique using Klenow fragment of *E. coli* Polymerase I, where the reaction proceeded for two minutes at 25° C., after which the steps of heating to 100° C. for two minutes, cooling, adding Klenow fragment, and reacting were repeated nine times. These results confirm the improved specificity of the amplification reaction employing the thermostable enzyme herein.

In a later cloning experiment with GM2064 and the contaminated primers, 43 out of 510 colorless plaques

(8%) had the (+)-strand insert. This suggests that approximately one-half of the 119 clones from Molt 4 contain the contaminant sequence.

Ten of the (+)-strand clones from Molt 4 were sequenced. Five were normal wild-type sequence and five had a single C to T mutation in the third position of the second codon of the gene (CAC to CAT). Four of the contaminant clones from GM2064 were sequenced and all four were normal.

Restriction site-modified primers may also be used to amplify and clone and partially sequence the human N-ras oncogene and to clone base pair segments of the HLA DQ-α, DQ-β and DR-β genes using the above technique.

Again, if the concentrations of $MgCl_2$ and nucleotides are reduced to 2 mM and 150–200 μM, respectively, and the minimum cycling temperature is increased from 37° C. to 45°–58° C., the specificity and efficiency of the amplification reaction can be increased.

EXAMPLE IV

Gene Retrieval

A. Identification of a DNA Sequence Probe for the TAQ Polymerase Gene

A specific DNA sequence probe for the Taq pol gene was obtained following immunological screening of a λgt11 expression library. *T. aquaticus* DNA was digested to completion with AluI, ligated with EcoRI 12-mer linkers (CCGGAATTCCGG, New England Biolabs), digested with EcoRI and ligated with dephosphorylated, EcoRI-digested λgt11 DNA (Promega Biotech). The ligated DNA was packaged (Gigapack Plus, Stratagene) and transfected into *E. coli* K-12 strain Y1090 (provided by R. Young).

The initial library of 2×10^5 plaques was screened (Young, R. A., and R. W. Davis (1983) *Science*, 222: 778–782) with a 1: 2000 dilution of a rabbit polyclonal antiserum raised to purified Taq polymerase (see Examples I and XIII). Candidate plaques were replated at limiting dilution and rescreened until homogeneous (~3 cycles). Phage were purified from candidate plaques which failed to react with preimmune serum and reacted with immune serum.

Candidate phage were used to lysogenize *E. coli* K-12 strain Y1089 (R. Young). Lysogens were screened for the production of an IPTG inducible fusion protein (larger than β-galactosidase) which reacted with the Taq polymerase antiserum. Solid phase, size-fractionated fusion proteins were used to affinity purify epitope-specific antibodies from the total polyclonal antiserum (Goldstein, L. S. B., et al. (1986) *J. Cell Biol.* 102: 2076–2087).

The "fished", epitope-selected antibodies were used, in turn, in a Western analysis to identify which λgt11 phage candidates encoded DNA sequences uniquely specific to Taq polymerase. One λgt11 phage candidate, designated λgt: 1, specifically selected antibodies from the total rabbit polyclonal Taq polymerase antiserum which uniquely reacted with both purified Taq polymerase and crude extract fractions containing Taq polymerase. This phage, λgt: 1, was used for further study.

The ~115 bp EcoRI-adapted AluI fragment of *Thermus aquaticus* DNA was labeled (Maniatis et al., supra) to generate a Taq polymerase-specific probe. The probe was used in Southern analyses and to screen a *T. aquaticus* DNA random genomic library.

B. Construction and Screening of a *Thermus Aquaticus* Random Genomic Library

Lambda phage Charon 35 (Wilhelmine, A. M. et al., supra) was annealed and ligated via its cohesive ends, digested to completion with BamHI, and the annealed arms were purified from the "stuffer" fragments by potassium acetate density gradient ultracentrifugation (Maniatis, et al., supra). *T. aquaticus* DNA was partially digested with Sau3A and the 15-20 kb size fraction purified by sucrose density gradient ultracentrifugation. The random genomic library was constructed by ligating the target and vector DNA fragments at a 1:1 molar ratio. The DNA was packaged and transfected into *E. coli* K-12 strains LE392 or K802. A library of >20,000 initial phage containing >99% recominants was amplified on *E. coli* K-12 strain LE392.

The CH35 Taq genomic phage library was screened (Maniatis et al., supra) with the readiolabeled EcoRI insert of λgt11: 1. Specifically hybridizing candidate phage plaques were purified and further analyzed. One phage, designated Ch35: 4-2, released ≧ four *T. aquaticus* DNA fragments upon digestion with HindIII (~8.0, 4.5, 0.8, 0.58 kb)

The four HindIII *T. aquaticus* DNA fragments were ligated with HindIII digested plasmid BSM13+ (3.2 kb, Vector Cloning Systems, San Diego) and individually cloned following transformation of *E. coli* K-12 strain DG98.

The ~8.0 kb HindIII DNA fragment from CH35: 4-2was isolated in plasmid pFC82 (11.2 kb), while the 4.5 kb HindIII DNA fragment from CH35: 4-2 was isolated in plasmid pFC83 (7.7 kb).

E. coli strain DG98 harboring pFC82 was shown to contain a thermostable, high temperature DNA polymerase activity (Table 1). In addition, these cells synthesize a new ~60 kd molecular weight polypeptide which is immunologically related to Taq DNA polymerase.

The Taq polymerase coding region of the 8.0 kb HindIII DNA fragment was further localized to the lac-promoter proximal 2.68 kb HindIII to Asp718 portion of the 8.0 kb HindIII fragment. This region was subcloned to yield plasmid pFC85 (6.0 kb). Upon induction with IPTG, *E. coli* DG98 cells harboring plasmid pFC85 synthesize up to 100-fold more thermostable, Taq polymerase-related activity (Table 1) than the original parent clone (pFC82/DG98). While cells harboring pFC85 synthesize a significant amount of a thermostable DNA polymerase activity, only a portion of the Taq pol DNA sequence is translated, resulting in the accumulation of a ~60 kd Taq polymerase-related polypeptide.

TABLE 1

Expression of a Thermostable DNA Polymerase Activity in *E. coli*#

Sample	Units•/ml IPTG	Units•/ml + IPTG
BSM13/DG98	—	0.02
pFC82/DG98	2.2	2.7
pFC85/DG98	11.9	643.8

#Cells were grown to late log phase (+/− IPTG, 10 mM), harvested, sonicated, heated at 75° C. for 20 minutes, centrifuged and the clarified supernatant assayed at 70° C. for DNA polymerase activity.

•1 unit = 1 nMole dCTP incorporated in 30 minutes.

EXAMPLE V

Expression of Taq Polymerase

The thermostable gene of the present invention can be expressed in any of a variety of bacterial expression vectors including DG141 (ATCC 39588) and $pP_LN_{RBS}ATG$, vectors disclosed in U.S. Pat. No. 4,711,845, the disclosure of which is incorporated herein by reference. Both of these host vectors are pBR322 derivatives that have either a sequence containing a tryptophan promoter-operator and ribosome binding site with an operably linked ATG start codon (DG141) or a sequence containing the lambda P_L promoter and gene N ribosome binding site operably linked to an ATG start codon ($pP_LN_{RBS}ATG$). Either one of these host vectors may be restricted with SacI, and blunt ended with Klenow or Sl nuclease to construct a convenient restriction site for subsequent insertion of the Taq polymerase gene.

The full-length Taq polymerase gene was constructed from the DNA insert fragments subcloned into plasmids pFC83 and pFC85 as follows. Vector BSM13+ (commercially available from Vector Cloning Systems, San Diego, Calif.) was digested at the unique HindIII site, repaired with Klenow and dNTPs, and ligated with T4 DNA ligase to a BglII octanucleotide linker, 5'-CAGATCTG-3' (New England Biolabs), and transformed into *E. coli* strain DG98. Plasmids were isolated from Amp^R $lacZ\alpha^+$ transformants. One of the clones was digested with BglII and Asp718 restriction enzymes, and the large vector fragment purified by gel electrophoresis.

Next, plasmid pFC83 was digested with BglII and HindIII and the ~730 base pair fragment was isolated. Plasmid pFC85 was digested with HindIII and Asp718 and the ~2.68 kb fragment isolated and joined in a three-piece ligation to the ~730 base pair BglII-HindIII fragment from pFC83 and the BglII-Asp718 vector fragment of BSM13+. This ligation mixture was used to transform *E. coli* strain DG98 (ATCC 39,768 deposited July 13, 1984) from which Amp^R colonies were selected and an ~6.58 kilobase plasmid (pLSG1) was isolated. Isopropyl-β-D-thiogalactoside (IPTG)-induced DG98 cells harboring pLSG1 synthesized Taq DNA polymerase indistinguishable in size from the native enzyme isolated from *T. aquaticus.*

Oligonucleotide-directed mutagenesis (see Zoller and Smith, *Nuc. Acids Res.* (1982) 10: 6487–6500) was used to simultaneously 1) introduce an SphI site within codons 3 to 5 of the Taq DNA polymerase gene sequence (see FIG. 1, nt 8–13), 2) increase the A/T content of four of the first seven condons without effecting a change in the encoded amino acids (within codons 2–7 in FIG. 1), 3) delete 170 nucleotides of the lacZ DNA and *T. aquaticus* DNA 5' to the DNA polymerase gene initiation codon.

Bacteriophage R408 (Russel, M., et al., *Gene,* (1986) 45: 333–338) was used to infect pLSG1/DG98 cells and direct the synthesis of the single-stranded DNA (ss) form (plus strand) of pLSG1. Purified pLSG1 ssDNA was annealed with purified PvuII-digested BSM13+ BglII vector fragments and the 47-mer mutagenic oligonucleotide DG26 (5'-CCCTTGGGCTCAAAAAGTGGAAGCATGCCTCTCATAGCTGTTTCCTG). Following extension with *E. coli* DNA polymerase I Klenow fragment, transformation of DG98 cells, and selection of Amp^R transformants, the colonies were screened with 5' ^{32}P-labeled DG26. Hybridizing candidates were screened for loss of the BglII restriction site, deletion of approximately 170 base pairs of lacZ:*T. aquaticus* DNA, and introduction of a unique SphI site. One candidate, designated pLSG2, was sequenced and shown to encode the desired sequence.

pLSG1 sequence:

S.D. 47bp BglII 105bp
CAGGAAACAGCT ATG [ACC ATG . . . AGATCT . . . AAC ATG]AGG GGG ATG CTG CCC CTC TTT

pLSG2 sequence:

S.D. Sph I
CAGGAAACAGCTATG AGA GGC ATG CTT CCA CTT TTT

Oligonucleotide-directed mutagenesis was used to introduce a unique BglII site in plasmid pLSG2 immediately following the TGA stop codon for the Taq polymerase gene (following nucleotide 2499 in FIG. 1). As above, bacteriophage R408 was used to generate the single-stranded (plus) form of plasmid pLSG2. Purified pLSG2 ssDNA was annealed with purified PvuII-digested BSM13+ BglII vector fragment and the 29-mer mutagenic oligonucleotide SC107 (5'-GCATGGGGTGGTAGATCTCACTCCTTGGC). Following extension with Klenow fragment (50 mM each dNTP), transformation of DG98 cells and selection for Amp^R transformants, colonies were screened with 5' ^{32}P-labeled SC107. Hybridizing candidates were screened for acquisition of a unique BglII site. One candidate, designated pSYC1578, was sequenced and shown to contain the desired sequence.

pLSG2 sequence:

. . . GCC AAG GAG TGA TAC CAC CCC ATG C . . .

pSYC1578 sequence:

BglII
. . . GCC AAG GAG TGA GATC TAC CAC CCC ATG C . . .

EXAMPLE VI

Construction of expression vectors pDG160 and pDG161

The Amp^R or Tet^R λP_L promoter, gene N ribosome binding site, polylinker, BT cry PRE (BT) (positive retroregulatory element, described in U.S. Pat. No. 4,666,848, issued May 19, 1987), in a ColE1 cop^{ts} vector were constructed from previously described plasmids and the duplex synthetic oligonucleotide linkers DG31 and DG32. The DG31/32 duplex linker encodes a 5' HindIII cohesive end followed by SacI, NcoI, KpnI-/Asp718, XmaI/SmaI recognition sites and a 3' BamHI cohesive end.

Asp718
SacI NcoI XmaI
DG31 5' AGCTTATGAGCTCCATGGTACCCCGGG
ATACTCGAGGTACCATGGGGCCCCTAG-5' DG32

A. Construction of Amp^R plasmid pDG160

Plasmid pFC54.t, a 5.96 kb plasmid described in U.S. Pat. No. 4,666,848, supra, was digested with HindIII and BamHI and the isolated vector fragment was ligated with a 5-fold molar excess of nonphosphorylated and annealed DG31/32 duplex. Following ligation, the DNA was digested with XbaI (to inactivate the parent vector IL-2 DNA fragment) and used to transform *E. coli* K12 strain DG116 to ampicillin resistance. Colonies were screened for loss of the des-ala-ser^{125} IL-2 mutein sequence and acquisition of the DG31/32 polylinker sequence by restriction enzyme digestion. The polylinker region in one candidate, designated pDG160, was sequenced and shown to encode the desired polylinker DNA sequence.

B. Construction of Tet^R plasmid pDG161

Plasmid pAW740CHB (ATCC 67,605), the source of a modified tetracycline resistance gene wherein the BamHI and HindIII restriction sites were eliminated, and which contains the λP_L promoter, gene N ribosome binding site, cry PRE in a ColE1 cop^{ts} vector, was digested to completion with HindIII and BamHI and the 4.19 kb vector fragment purified by agarose gel electrophoresis. The purified vector DNA fragment was ligated with a 5-fold molar excess of nonphosphorylated annealed DG31/32 duplex. *E. coli* K12 strain DG116 was transformed with a portion of the DNA, and Tet^R colonies screened for presence of 4.2 kb plasmids. Several candidates were further screened by restriction enzyme digestion and the polylinker region sequenced by the Sanger method. One of the candidates with the desired sequence was designated pDG161.

EXAMPLE VII

A. Construction of an Amp^R P_L promoter, gene N ribosome binding site, (N_{RBS}) Taq polymerase (832) BT cry PRE, cop^{ts} expression vector

To express the full-length (832 amino acid) mutated Taq polymerase sequence encoded by plasmid pSYC1578 under the control of the λP_L promoter and gene N ribosome binding site, we used plasmids pSYC1578 and pFC54.t. Plasmid pSYC1578 was digested with SphI and BglII and the resulting approximate 2.5 kb Taq polymerase gene fragment purified by agarose gel electrophoresis and electroelution. Plasmid pFC54.t was digested to completion with HindIII and BamHI and the vector fragment purified by agarose gel electrophoresis. The synthetic oligonucleotides DG27 (5'-AGCTTATGAGAGGCATG) and DG28 (5'-CCTCTCATA) were synthesized and annealed. Purified pFC54.t fragment (0.085 pmoles), purified Taq polymerase gene fragment (0.25 pmoles) and annealed nonphosphorylated DG27/28 duplex adaptor (0.43 pmoles) were combined in 30 μl and ligated at 14° C. A portion of the ligated DNA was heated to 75° C. (15 minutes) to inactivate the DNA ligase in the samples and treated with XbaI to linearize (inactivate) any IL-2 mutein containing ligation products. The ligated and digested DNA (approximately 100 ng) was used to transform *E. coli* K12 strain DG116 to ampicillin resistance. Amp^R colonies were screened for the presence of an approximate 8 kb plasmid which yielded the expected digestion products with HindIII (621 bp+7,410 bp), EcoRI (3,250 bp+4,781 bp) and SphI (8,031 bp), Asp718 (8,031 bp), BamHI (8,031 bp) and PvuII (4,090 bp+3,477 bp+464 bp). Several candidates were subjected to DNA sequence analysis at the 5' λP_L:TaqPol junction and the 3' TaqPol:BT junction. One of the candidates was also screened with an anti-Taq polymerase antibody for the synthesis of an approximate 90 kd immunoreactive antigen. Single colonies were transferred from a 30° C. culture plate to a 41° C. culture plate for two hours. The colonies were scraped with a toothpick from both the 30° C. and 41° C. plates, boiled in SDS loading buffer, subjected to SDS-PAGE electrophoresis and the separated proteins transferred to a nitrocellulose membrane. The membranes were probed with a 1:6,000 dilution of a polyclonal anti-Taq antibody and developed with a goat anti-rabbit HRP conjugate. All of the candidates tested showed evidence of temperature inducible approximate 90 kd Taq polymerase-related protein. One of the several plasmid candidates which directed the synthesis of Taq polymerase in *E. coli* and contained the expected DNA sequence was designated pLSG5.

B. Construction of a Tet^R P_L promoter, gene N ribosome binding site, Taq polymerase (832) BT cry PRE cop^{ts} expression vector

To express the full length (832 amino acid) mutated Taq polymerase sequence encoded by plasmid pSYC1578 under control of the λP_L promoter and gene N ribosome binding site in a Tet^R vector, we used plasmids pSYC1578 and pAW740CHB. Plasmid pSYC1578 was digested with SphI and BglII and the resulting approximate 2.5 kb Taq polymerase gene fragment was purified by agarose gel electrophoresis and electroelution. Plasmid pAW740CHB was digested to completion with HindIII and BamHI and the resulting 4.19 kb vector fragment purified by agarose gel electrophoresis and electroelution. The synthetic oligonucleotides DG27 and DG28 (described previously) were annealed. Purified pAW740CHB vector fragment (0.12 pmoles) was ligated with purified Taq polymerase gene fragment (0.24 pmoles) and annealed nonphosphorylated DG27/28 duplex adaptor (0.24 pmoles) in 30 μl at 14° C. A portion of the ligated DNA (100 ng) was used to transform *E. coli* K12 strain DG116 to tetracycline resistance. Tet^R candidates were screened for the presence of an approximate 6.7 kb plasmid which yielded the expected digestion products with HindIII (621 bp+6,074 bp), EcoRI (3,445 bp+3,250 bp), Asp718 (6,695 bp), SphI (3,445 bp+3,250 bp), BamHI (6,695 bp) and PvuII (3,477 bp+2,754 bp+464 bp). Several candidates were subjected to DNA sequence analysis at the 5' λP_L:TaqPol junction and the 3' TaqPol:BT junction. Candidates were also screened by single colony immunoblot as described above for the temperature inducible synthesis of Taq polymerase. One of the plasmid candidates which directed the synthesis of Taq polymerase in *E. coli* and contained the expected DNA sequence was designated pLSG6.

Case 3:11-cv-00056-IEG -WVG Document 1-2 Filed 01/10/11 Page 31 of 90

EXAMPLE VIII

Construction of a Met4 (v3) 829 Amino Acid Form of Taq Polymerase

The predicted fourth codon of native Taq polymerase directs the incorporation of a methionine residue (see pLSG1 and pLSG2 5′ sequences above). To obtain a further mutated form of the Taq polymerase gene that would direct the synthesis of an 829 amino acid primary translation product we used plasmids pSYC1578 and pDG161. Plasmid pSYC1578 was digested with SphI, treated with *E. coli* DNA polymerase I Klenow fragment in the presence of dGTP to remove the four-base 3′ cohesive end and generate a CTT (leucine, 5th codon) blunt end. Following inactivation of the DNA polymerase and concentration of the sample, the DNA was digested with BglII and the approximate 2.5 kb Taq polymerase gene fragment purified by agarose gel electrophoresis and electroelution. Plasmid pDG161 was digested to completion with SacI, repaired with *E. coli* DNA polymerase I Klenow fragment in the presence of dGTP to remove the four base 3′ cohesive end and generate an ATG terminated duplex blunt end. Following inactivation of the polymerase, the sample was digested with BamHI.

Digested pDG161 (0.146 pmole) and purified Taq polymerase fragment (0.295 pmole) were ligated at 30 μg/ml under sticky end conditions overnight. The partially ligated DNA sample (BamHI/BglII ends) was diluted to 15 μg/ml and ligated for five hours under blunt end conditions. The DNA ligase was inactivated (75° C., 10 minutes) and the sample digested with NcoI to linearize any ligation products containing the pDG161 polylinker sequence. Sixty nanograms of the ligated and digested DNA was used to transform *E. coli* K12 strain DG116 to tetracycline resistance. Tet^R candidates were screened for the presence of an approximate 6.7 kb plasmid which yielded the expected digestion products when treated with HindIII (612 bp+6,074 bp), EcoRI (3,445 bp+3,241 bp) and SphI (6,686 bp). Colonies were screened as above by single colony immunoblot for the temperature inducible synthesis of an approximate 90 kd Taq polymerase-related polypeptide. One of the plasmids, designated pLSG7, that directed the synthesis of a Taq polymerase-related polypeptide was subjected to Sanger sequence determination at the 5′ λP_L promoter:Taq polymerase junction and the 3′ Taq polymerase:BT junction. Analysis of the DNA sequence at the 5′ junction confirmed the restriction enzyme analysis (loss of one of the SphI sites and a 612 bp HindIII fragment, slightly smaller than the 621 bp HindIII fragment in pLSG6) and indicated the derivation of a plasmid encoding an 829 amino acid form of Taq polymerase.

EXAMPLE IX

Construction of Met289 (v289) 544 Amino Acid Form of Taq Polymerase

During purification of native Taq polymerase (Example XIII) we obtained an altered form of Taq polymerase that catalyzed the template dependent incorporation of dNTP at 70° C. This altered form of Taq polymerase was immunologically related to the approximate 90 kd form described in Example XIII but was of lower molecular weight. Based on mobility, relative to BSA and ovalbumin following SDS-PAGE electrophoresis, the apparent molecular weight of this form is approximately 61 kd. This altered form of the enzyme is not present in carefully prepared crude extracts of *Thermus aquaticus* cells as determined by SDS-PAGE Western blot analysis or in situ DNA polymerase activity determination (Spanos, A., and Hubscher, U. (1983) *Meth. Enz.* 91: 263–277) following SDS-PAGE gel electrophoresis. This form appears to be proteolytic artifact that may arise during sample handling. This lower molecular weight form was purified to homogeneity and subjected to N-terminal sequence determination on an ABI automated gas phase sequencer. Comparison of the obtained N-terminal sequence with the predicted amino acid sequence of the Taq polymerase gene (see FIG. 1) indicates this shorter form arose as a result of proteolytic cleavage between glu_{289} and ser_{290}.

To obtain a further truncated form of a Taq polymerase gene that would direct the synthesis of a 544 amino acid primary translation product we used plasmids pFC54.t, pSYC1578 and the complementary synthetic oligonucleotides DG29 (5′-AGCTTATGTCTCCAAAAGCT) and DG30 (5′-AGCTTTTGGAGACATA). Plasmid pFC54.t was digested to completion with HindIII and BamHI. Plasmid pSYC1578 was digested with BstXI and treated with *E. coli* DNA polymerase I Klenow fragment in the presence of all 4 dNTPs to remove the 4 nucleotide 3′ cohesive end and generate a CTG-terminated duplex blunt end encoding leu_{294} in the Taq polymerase sequence (see pLSG1, nucleotide 880). The DNA sample was digested to completion with BglII and the approximate 1.6 kb BstXI (repaired)/BglII Taq DNA fragment was purified by agarose gel electrophoresis and electroelution. The pFC54.t plasmid digest (0.1 pmole) was ligated with the Taq polymerase gene fragment (0.3 pmole) and annealed nonphosphorylated DG29/DG30 duplex adaptor (0.5 pmole) under sticky ligase conditions at 30 μg/ml, 15° C. overnight. The DNA was diluted to approximately 10 microgram per ml and ligation continued under blunt end conditions. The ligated DNA sample was digested with XbaI to linearize (inactivate) any IL-2 mutein-encoding ligation products. 80 nanograms of the ligated and digested DNA was used to transform *E. coli* K12 strain DG116 to ampicillin resistance. Amp^R candidates were screened for the presence of an approximate 7.17 kb plasmid which yielded the expected digestion products with EcoRI (4,781 bp+2,386 bp), PstI (4,138 bp+3,029 bp), ApaI (7,167 bp) and HindIII/PstI (3,400 bp+3,029 bp+738 bp). *E. coli* colonies harboring candidate plasmids were screened as above by single colony immunoblot for the temperature-inducible synthesis of an approximate 61 kd Taq polymerase related polypeptide. In addition, candidate plasmids were subjected to DNA sequence determination at the 5′ λP_L promoter:Taq DNA junction and the 3′ Taq DNA:BT cry PRE junction. One of the plasmids encoding the intended DNA sequence and directing the synthesis of a temperature-inducible 61 kd Taq polymerase related polypeptide was designated pLSG8.

Yet another truncated Taq polymerase gene contained within the ~2.68 kb HindIII-Asp718 fragment of plasmid pFC85 can be expressed using, for example, plasmid $pP_LN_{RBS}ATG$, by operably linking the amino-terminal HindIII restriction site encoding the Taq pol gene to an ATG initiation codon. The product of this fusion upon expression will yield an ~70,000–72,000 dalton truncated polymerase.

This specific construction can be made by digesting plasmid pFC85 with HindIII and treating with Klenow fragment in the presence of dATP and dGTP. The resulting fragment is treated further with Sl nuclease to remove any single-stranded extensions and the resulting DNA digested with Asp718 and treated with Klenow fragment in the presence of all four dNTPs. The recovered fragment can be ligated using T4 DNA ligase to dephosphorylated plasmid $pP_LN_{RBS}ATG$, which had been digested with SacI and treated with Klenow fragment in the presence of dGTP to construct an ATG blunt end. This ligation mixture can then be used to transform *E. coli* DG116 and the transformants screened for production of Taq polymerase. Expression can be confirmed by Western immunoblot analysis and activity analysis.

EXAMPLE X

Construction of Amp^R Trp Promoter Operator, TrpL Ribosome Binding Site, Taq Polymerase (832) BT Cry PRE Cop^{ts} Expression Vector

To substitute the *E. coli* trp operon promoter/operator and leader peptide ribosome binding site, we used plasmids pLSG5 and pFC52. pFC52 was the source of the trp promoter, cop^{ts} and ampicillin resistant determinants. However, plasmid pCS4, described in U.S. Pat. No. 4,711,845, supra, the disclosure of which is incorporated herein by reference, may be used to provide the identical fragment. Plasmid pLSG5 was digested to completion with SphI. The SphI was inactivated (70° C., 10 minutes) and the digested DNA was ligated overnight at 15° C. with an excess of annealed nonphosphorylated DG27/28 duplex adaptor (see above). The T4 DNA ligase was inactivated (70° C., 10 minutes) and the DNA digested to completion with MluI. The DNA sample was sequentially extracted with phenol and ether, ethanol precipitated and finally resuspended in 10 mM Tris chloride pH 8, 1 mM EDTA. Plasmid pFC52 (or pCS4) was digested to completion with MluI and extracted with phenol, ether and concentrated as above. The DNA sample was digested to completion with HindIII and the HindIII inactivated (75° C., 15 minutes). The pLSG5 and pFC52 samples were ligated overnight in equal molar ratio and at 30 μg/ml under sticky end conditions. The T4 ligase was inactivated (70° C., 10 minutes) and the ligated DNA was digested with XbaI to linearize (inactivate) any IL-2 encoding ligation products (from the pFC52 unwanted, 1.65 kb HindIII/MluI DNA fragment). *E. coli* K12 strain DG116 was transformed to ampicillin resistance with 30 nanogram of the ligated DNA. Amp^R colonies were screened for the presence of approximate 7.78 kb plasmids which yielded the expected digestion products with EcoRI (4,781 bp+3,002 bp), SphI (7,783 bp), HindIII (7,162 bp+621 bp), ClaI (7,783 bp) and ClaI/MluI (3,905 bp+3,878 bp). Candidate colonies were further screened for expression of an approximate 90 kd Taq polymerase related protein by single colony SDS-PAGE immunoblotting (as above). Plasmids from two of the candidates showing the intended properties were transformed into *E. coli* K12 strain KB2 (ATCC No. 53075).

By Western immunoblot, both plasmids in both hosts were shown to direct the synthesis of an approximate 90 kd Taq polymerase-related polypeptide upon trp limitation. By Comassie staining of SDS-PAGE fractionated whole cell extract proteins, the trp promoter/Taq polymerase plasmids in *E. coli* K12 strain KB2 direct the accumulation of significantly more Taq polymerase than in *E. coli* K12 strain DG116. One of the plasmids was designated pLSG10.

EXAMPLE XI

Synthesis of Recombinant Taq DNA Polymerase Activity in *E. coli*

E. coli K12 (DG116) strains harboring plasmids pDG160, or pLSG5, or pLSG6 were grown at 32° C. in Bonner-Vogel minimal salts media containing 0.5% glucose, 10 μg/ml thiamine, 0.25% (w/v) Difco casamino acids and ampicillin (100 μg/ml) or tetracycline (10 μg/ml) as appropriate. Cells were grown to A_{600} of about 0.8 and shifted to 37° C. to simultaneously derepress the lambda P_L promoter (inactivation of cI_{857} repressor) and increase the copy number of the ColE1 cop^{ts} plasmid vector. After six-nine hours of growth at 37° C., aliquots of the cells were harvested, the cells centrifuged and the pellets stored at −70° C.

Alternatively, *E. coli* K12 strain KB2 harboring plasmid pLSG10 was grown for eight hours at 32° C. in Bonner-Vogel minimal salts media containing 0.5% glucose, 5 μg/ml tryptophan, 10 μ/ml thiamine, 0.25% Difco casamino acids and 100 μg/ml ampicillin to an A_{600} of 3.0. Cells were harvested as above.

Cell pellets were resuspended to about 62.5 A_{600}/ml (~150–160 μg total protein/ml) in 50 mM Tris-Cl, pH 7.5, 1 mM EDTA, 2.4 mM PMSF and 0.5 μg/ml leupeptin and lysed by sonication. Aliquots of the sonicated extracts were subjected to SDS-PAGE and analyzed by Coomassie staining and Western immunoblotting with rabbit polyclonal anti-Taq polymerase antibody. In addition, portions of the extracts were assayed in a high temperature (74° C.) DNA polymerase assay (see Example XIII below).

Western immunoblotting showed significant induction and synthesis of an approximately 94 kd Taq DNA polymerase related polypeptide in induced strains harboring plasmids pLSG5, 6, and 10. Coomassie blue staining of SDS-PAGE-separated total cell protein revaled the presence of a new predominant protein at ~94 kd in these induced strains. Finally, high temperature activity assays confirmed the significant level of recombinant Taq DNA polymerase synthesis in these *E. coli* strains (see table, below).

Plasmid Host	Taq Pol Gene	Promoter	Uninduced (−) or Induced (+)	Units*/ OD_{600}
pDG160/DG116	−	P_L	− or +	<1.0
pLSG5/DG116	+	P_L	−	23
pLSG5/DG116	+	P_L	+	308
pLSG6/DG116	+	P_L	−	5
pLSG6/DG116	+	P_L	+	170
pLSG10/KB2	+	Trp	+	300

*1 unit = 10 nmole total nucleotide incorporated at 74° C./30 minutes.

EXAMPLE XII

Purification of Recombinant Taq DNA Polymerase

E. coli strain DG116 harboring plasmid pLSG5 was grown in a 10 L fermentor. The medium was 10 mM $(NH_4)_2SO_4$, 25 mM KH_2PO_4, 4 mM $Na_3Citrate$, 400 μM $FeCl_3$, 28 μM $ZnCl_2$, 34 μM $CoCl_2$, 33 μM $NaMoO_4$, 27 μM $CaCl_2$, 30 μM $CuCl_2$, and 32 μM H_3BO_3. The medium was adjusted to pH 6.5 with NaOH, ~15 mM, and sterilized. The following sterile components

were added: 20 mg/l thiamine.HCl, 3 mM $MgSO_4$, 10 g/l glucose and 12.5 mg/l ampicillin. The pH was adjusted to 6.8 and held there using NH_4OH. Glucose was fed to the culture in conjunction with the alkali demand, to maintain a glucose concentration at 40% of air saturation, by automatic increases in rpm (350 to 1000) and airflow (2 to 5 l/min). Foaming was controlled on demand using polypropylene glycol.

The fermentor was inoculated with cells and grown to A_{680}=5.0 (14.25 hours). The temperature was raised to 37° C. to induce synthesis of recombinant Taq polymerase and growth continued for five hours to A_{680} of 16.5.

Unless otherwise indicated, all purification steps were conducted at 4° C. Twenty grams (wet weight) of induced frozen *E. coli* K12 strain DG116 harboring plasmid pLSG5 was thawed in 3 volumes of 50 mM Tris-Cl, pH 7.5, 1 mM EDTA, 3 mM PMSF, 0.64 μg/ml leupeptin and disrupted in a French Press at 20,000 psi. The lysate was adjusted to 5.5× cell volume with additional buffer and sonicated (4×30 seconds) to reduce viscosity (Fraction I). The crude total cell lysate was adjusted to 0.2M $(NH_4)_2SO_4$ (26.43 g/l) and centrifuged for 15 minutes at 20,000×G. The supernatant (Fraction II) was heated to 75° C. (in a 100° C. water bath) and maintained at 72°–75° C. for 15 minutes to denature *E. coli* host proteins. The sample was rapidly cooled to 4° C. by swirling in an ice water bath. After 20 minutes at 0° C., the sample was centrifuged at 20,000×G for 15 minutes to precipitate the denatured proteins. The supernatant (Fraction III) was applied at 4 ml/hr to a 6 ml Phenyl-Sepharose CL-4B (Pharmacia) column equilibrated with 50 mM Tris-Cl, pH 7.5, 1 mM EDTA (Buffer A) containing 0.2M $(NH_4)_2SO_4$. The column was sequentially washed with 3–10 column volumes of a) the same buffer, b) Buffer A, c) Buffer A containing 20% ethylene glycol to remove nucleic acids and non-Taq polymerase proteins. Taq DNA polymerase activity was eluted with 60 ml linear gradient of 0–4M urea in Buffer A containing 20% ethylene glycol. The active fractions (~2M urea) were pooled (Fraction IV) and applied at 3 ml/hr to a 12 ml (1.5×6.0 cm) Heparin-Sepharose CL-6B (Pharmacia) column equilibrated in 50 mM Tris-Cl, pH 7.5, 0.1 mM EDTA, 0.2% Tween 20 (Buffer B) containing 0.1M KCl. The column was washed with 2 column volumes of Buffer B containing 0.15M KCl. The Taq polymerase was eluted with a 120 ml linear gradient of 0.15–0.65M KCl in Buffer B. The Taq polymerase eluted as a single A_{280} and activity peak at ~0.29M KCl.

Purified recombinant and native Taq polymerase proteins comigrate following electrophoresis on SDS-PAGE and staining with Coomassie blue. The purified Taq polymerase proteins migrate slightly faster than purified Phosphorylase B (Pharmacia), consistent with a molecular weight predicted from the DNA sequence (of pLSG5) of 93,920 daltons.

The peak activity fractions were pooled and a portion subjected to N-terminal amino acid sequence determination on an Applied Biosystems gas phase sequencer. In contrast to native Taq polymerase which has a blocked amino terminus, the sequence of the purified recombinant Taq polymerase and the individual cycle yields were consistent with the sequence predicted for the amino terminus of the Taq polymerase protein encoded by plasmid pLSG5.

The recombinant Taq polymerase encoded by plasmid pLSG5 and purified as described could amplify a human "single copy" sequence. Using a low temperature limit of 55° C., extension temperature of 72° C., upper temperature limit of 94° C. and a 2–2.5 minute cycle time, comparable yields and efficiency were noted for native and recombinant Taq polymerase using 1–2 units/100 μl PCR.

EXAMPLE XIII

Purification

The thermostable polymerase may be purified directly from a culture of *Thermus aquaticus* following the example disclosed below or, alternatively, from a bacterial culture containing the recombinantly produced enzyme with only minor modifications necessary in the preparation of the crude extract.

After harvesting by centrifugation, 60 grams of cells were resuspended in 75 ml of a buffer consisting of 50 mM Tris-Cl pH 8, 1 mM EDTA. Cells were lysed in a French Press at 14,000–16,000 PSI after which 4 volumes (300 ml) of additional Tris-EDTA were added. Buffer A (β-mercaptoethanol to 5 mM and NP-40 and Tween 20 to 0.5% (v/v) each) was added and the solution was sonicated thoroughly while cooling. The resultant homogeneous suspension was diluted further with Buffer A such that the final volume was 7.5–8 times the starting cell weight; this was designated Fraction I.

The polymerase activity in Fraction I and subsequent fractions was determined in a 50 μl mixture containing 0.025M TAPS-HCl pH 9.4 (20° C.), 0.002M $MgCl_2$, 0.05M KCl, 1 mM 2-mercaptoethanol, 0.2 mM each dGTP, dATP, TTP, 0.1 mM dCTP [α-^{32}P, 0.05 Ci/mM], 12.5 μg "activated" salmon sperm DNA and 0.01–0.2 units of the polymerase (diluted in 10 mM Tris-HCl, pH 8, 50 mM KCl, 1 mg/ml autoclaved gelatin, 0.5% NP-40, 0.5% Tween 20, and 1 mM 2-mercaptoethanol). One unit corresponds to 10 nmoles of product synthesized in 30 minutes. "Activated" DNA is a native preparation of DNA after partial hydrolysis with DNase I until 5% of the DNA was transferred to the acid-soluble fraction. The reaction was conducted at 74° C. for 10 minutes and then 40 μl was transferred to 1.0 ml of 50 μg/ml carrier DNA in 2 mM EDTA at 0° C. An equal volume (1.0 ml) of 20% TCA, 2% sodium pyrophosphate was added. After 15–20 minutes at 0° C. the samples were filtered through Whatman GF/C discs and extensively washed with cold 5% TCA-1% pyrophosphate, followed by cold 95% ethanol, dried and counted.

Fraction I was centrifuged for two hours at 35,000 rpm in a Beckman TI 45 rotor at 2° C. and the collected supernatant was designated Fraction II.

The Taq polymerase activity was precipitated with Polymin P (BRL, Gaithersburg, MD) (10%, w/v, adjusted to pH 7.5 and autoclaved) after the minimum amount of Polymin P necessary to precipitate 90–95% of the activity was determined, which amount was generally found to be between 0.25% and 0.3% final volume.

An appropriate level of Polymin P was added slowly to Fraction II while stirring for 15 minutes at 0° C. This solution was centrifuged at 13,000 rpm for 20 minutes in a Beckman JA 14 rotor at 2° C. The supernatant was assayed for activity and the pellet was resuspended in 1/5 volume of 0.5×Buffer A (diluted 1:2 with H_2O). This suspension was recentrifuged and the pellet resuspended in ¼ volume of Buffer A containing 0.4M KCl. This suspension was homogenized thoroughly and left

overnight at 4° C. The homogenate was centrifuged as above and the collected supernatant designated Fraction III.

The protein fraction was collected by "precipitation" at 75% saturation of ammonium sulfate, centrifuged (at 27,000 rpm, SW27 rotor, 30 minutes) and the floating pellicle was resuspended in 50 mM Tris-Cl pH 8, 1 mM EDTA. These steps were repeated and the protein suspension was dialyzed extensively with P-cell buffer (20 mM KPO_4 pH 7.5, 0.5 mM EDTA, 5 mM β-mercaptoethanol, 5% (w/v) glycerol, 0.5% (v/v) NP-40 and Tween 20) containing 80 mM KCl.

The dialysate was transferred to a centrifuge bottle to which was added any recovered protein from sacks rinsed with the P-cell buffer containing 80 mM KCl. Centrifugation was performed at 20,000×g and the time was reduced to 15 minutes. The supernatant was saved and any pellet remaining was washed, extracted with P-cell buffer and 80 mM KCl, and recentrifuged. The supernatants were then combined to form Fraction IV.

Fraction IV was applied to a 2.2×22-cm column of phosphocellulose, equilibrated with the P-cell buffer containing 80 mM KCl. The column was washed (2.5–3 column volumes) with the same buffer and the protein eluted using a linear gradient of 80 to 400 mM KCl in P-cell buffer. Fractions containing DNA polymerase activity (~0.18–0.20M KCl) were pooled and concentrated 3–4 fold on an Amicon stirred cell and YM30 membrane. The cell was rinsed with the P-cell buffer without KCl and added to the fraction concentrate (0.15M KCl adjusted final volume) to form Fraction V.

Fraction V was applied to a 5 ml Heparin Sepharose CL-6B column (Pharmacia) equilibrated with P-cell buffer and 0.15M KCl. The column was washed with 0.15M KCl buffer (3–4 column volumes) and the protein eluted with a linear gradient from 0.15 to 0.65M KCl in P-cell buffer. A 1:10 dilution into diluent without gelatin was made for SDS-PAGE analysis and a subsequent 1:20 dilution into diluent with 1 mg/ml gelatin was made for use in enzyme assays. The activity fractions (eluting at ~0.3M KCl) were assayed on supercoiled DNA template for specific and non-specific endoncleases/topoisomerase by electrophoretically detecting the change in molecular weight of supercoiled plasmid DNA after incubation with an excess of DNA polymerase. Exonuclease contamination was detected following incubation with small linear DNA fragments. In peak fractions, an ~88–92 kd protein was found to be the major band. The major pool, designated Fraction VI, had the highest polymerase activity with minimal detectable endonuclease activity when this pool was assayed for 30 minutes at 55° C. with ~3–5 polymerase units/600 ng DNA.

Fraction VI was dialyzed against 10 mM KPO_4 pH 7.5, 5 mM β-mercaptoethanol, 5% glycerol, 0.2% NP-40, and 0.2% Tween 20 (HA buffer). The dialyzed sample was applied to a 3 ml column of hydroxyapatite and the enzyme eluted with a linear gradient of 10 to 250 mM KPO_4 pH 7.5, HA buffer. DNA polymerase activity began to elute at 75 mM KPO_4 with the peak at 100 mM KPO_4. Active peak fractions were assayed at 1:100–1:300 dilution. As in the prior chromatography step, a 1:10 dilution in diluent was prepared without gelatin for SDS-PAGE analysis. Fractions with no significant endonuclease or double-strand exonuclease when assayed at 55° C. with 5 polymerase units were pooled and designated Fraction VII.

Fraction VII was dialyzed against a solution of 25 mM sodium acetate pH 5.2, 5% glycerol, 5 mM β-mercaptoethanol, 0.1 mM EDTA, 0.1% NP-40, and 0.1% Tween 20, adjusted to pH 5 at room temperature. The dialyzed sample was applied to a 2 ml DEAE-Tris-Acryl-M (LKB) column pre-equilibrated and subsequently washed with the same buffer. The fraction containing polymerase activity that did not adhere to the column was pooled and adjusted to 50 mM NaCl in the same buffer to yield Fraction VIII.

Fraction VIII was applied to a 2 ml CM-Tris-Acryl M (LKB) column equilibrated with the same buffer (25 mM sodium acetate, 50 mM NaCl, 5% glycerol, 0.1 mM EDTA, 0.1% NP-40, and 0.1% Tween 20). The column was washed with 4–5 column volumes of the same buffer and the enzyme eluted with a linear gradient from 50 to 400 mM NaCl in sodium acetate buffer. The polymerase activity peak eluted ~0.15–0.20M NaCl. The polymerase activity was assayed at 1:300 to 1:500 dilution with the first dilution 1:10 into diluent without gelatin for the SDS-PAGE analysis. An assay across the activity peak on supercoiled DNA templates for specific and non-specific endonuclease/topoisomerase using DNA polymerase assay salts (25 mM TAPS-HCl pH 9.4, 2.0 mM $MgCl_2$ and 50 mM KCl) at 74° C. was performed, as well as assays for nucleases on M13 ss DNA and pBR322 fragments. Active fractions with no detectable nuclease(s) were pooled and run on a silver stained SDS-PAGE mini gel. The results show a single ~88–92 kd band with a specific activity of ~200,000 units/mg.

This specific activity is more than an order of magnitude higher than that claimed for the previously isolated Taq polymerase and is at least an order of magnitude higher than that for *E. coli* polymerase 1.

EXAMPLE XIV

The Taq polymerase purified as described above in Example XIII was found to be free of any contaminating Taq endonuclease and exonuclease activities. In addition, the Taq polymerase is preferably stored in storage buffer containing from about 0.1 to about 0.5% volume/volume of each non-ionic polymeric detergent employed. More preferably the storage buffer consists of 50% (v/v) glycerol, 100 mM KCl, 20 mM Tris-Cl pH 8.0, 0.1 mM ethylenediaminetetraacetic acid (EDTA), 1 mM dithiothreitol, 0.5% v/v NP-40, 0.5% v/v Tween 20, and 200 μg/ml gelatin, and is preferably stored at −20° C.

The stored Taq polymerase was diluted in a buffer consisting of 25 mM Tris Cl pH 8.0, 20 mM KCl, 1 mM β-mercaptoethanol, 0.5% NP-40, 0.5% Tween-20, and 500 μg/ml gelatin. A reaction buffer was then prepared containing 50 mM KCl, 10 mM Tris-Cl, pH 8.3, 1.5 mM $MgCl_2$, 0.01% (w/v) gelatin, 200 μM each dNTP, 1 μM each of the primers that define a 500 base pair target sequence on a control template from bacteriophage λ, and 2.0–2.5 units Taq polymerase/assay in a final volume of 100 μl. Template was added to the reaction buffer, the sample placed in a 0.5 ml polypropylene tube, and the sample topped with 100 μl of heavy white mineral oil to prevent evaporation.

At least a 10^5-fold amplification was achieved when the following conditions were employed, using 1 ng of control template (bacteriophage λDNA) where the target sequence represented approximately 1% of the starting mass of DNA.

First the template mixture was denatured for one minute, 30 seconds at 94° C. by placing the tube in a heat bath. Then the tube was placed in a heat bath at 37° C. for two minutes. Then the tube was placed in a heat bath at 72° C. for three minutes, and then in the heat bath at 94° C. for one minute. This cycle was repeated for a total of 25 cycles. At the end of the 25th cycle, the heat denaturation step at 94° C. was omitted and replaced by extending the 72° C. incubation step by an additional three minutes. Following termination of the assay, the samples were allowed to cool to room temperature and analyzed as described in previous examples.

The template may be optimally amplified with a different concentration of dNTPs and a different amount of Taq polymerase. Also, the size of the target sequence in the DNA sample will directly impact the minimum time required for proper extension (72° C. incubation step). An optimization of the temperature cycling profile should be performed for each individual template to be amplified, to obtain maximum efficiency.

EXAMPLE XV

Taq polymerase purified as described above in Example I was formulated for storage as described in the previous example, but without the non-ionic polymeric detergents. When assayed for activity as described in that example, the enzyme storage mixture was found to be inactive. When the NP-40 and Tween 20 were added to the storage buffer, the full enzyme activity was restored, indicating that the presence of the non-ionic detergents is necessary to the stability of the enzyme formulation.

EXAMPLE XVI

Several 1 μg samples of human genomic DNA were subjected to 20–35 cycles of amplification as described in Example II, with equivalent units of either Klenow fragment or Taq polymerase, and analyzed by agarose gel electrophoresis and Southern blot. The primers used in these reactions, PC03 and PC04, direct the synthesis ~of a 110-bp segment of the human beta-globin gene. The Klenow polymerase amplifications exhibited the smear of DNA typically observed with this enzyme, the apparent cause of which is the non-specific annealing and extension of primers to unrelated genomic sequences under what were essentially non-stringent hybridization conditions (1×Klenow salts at 37° C.). Nevertheless, by Southern blot a specific 110-bp beta-globin target fragment was detected in all lanes. A substantially different electrophoretic pattern was seen in the amplifications done with Taq polymerase where the single major band is the 110-bp target sequence. This remarkable specificity was undoubtedly due to the temperature at which the primers were extended.

Although, like Klenow fragment amplifications, the annealing step was performed at 37° C., the temperature of Taq-catalyzed reactions had to be raised to about 70° C. before the enzyme exhibited significant activity. During this transition from 37° to 70° C., poorly matched primer-template hybrids (which formed at 37° C.) disassociated so that by the time the reaction reached an enzyme-activating temperature, only highly complementary substrate was available for extension. This specificity also results in a greater yield of target sequence than similar amplifications done with Klenow fragment because the non-specific extension products effectively compete for the polymerase, thereby reducing the amount of 110-mer that can be made by the Klenow fragment.

EXAMPLE XVII

Amplification was carried out of a sample containing 1 μg Molt 4 DNA, 50 mM KCl, 10 mM Tris pH 8.3, 10 mM $MgCl_2$, 0.01% gelatin, 1 μM of each of the following primers (to amplify a 150 bp region):

5'-CATGCCTCTTTGCACCATTC-3'(RS79)

and

5'-TGGTAGCTGGATTGTAGCTG-3'(RS80)

1.5 mM of each dNTP, and 5.0 units of Taq polymerase per 100 μl reaction volume. Three additional samples were prepared containing 2.5, 1.3, or 0.6 units of Taq polymerase. The amplification was carried out in the temperature cycling machine described above using the following cycle, for 30 cycles:

from 70° to 98° C. for 1 minute
hold at 98° C. for 1 minute
from 98° C. to 35°, 45° or 55° C. for 1 minute
hold at 35°, 45° or 55° C. for 1 minute
from 35°, 45° or 55° C. to 70° C. for 1 minute
hold at 70° C. for 30 seconds

At 35° C. annealing temperature, the 2.5 units/100 μl Taq enzyme dilution gave the best-signal-to noise ratio by agarose gel electrophoresis over all other Taq polymerase concentrations. At 45° C., the 5 units/100 μl Taq enzyme gave the best signal-to-noise ratio over the other concentrations. At 55° C., the 5 units/100 μl Taq enzyme gave the best signal-to-noise ratio over the other concentrations and over the 45° C. annealing and improved yield. The Taq polymerase has more specificity and better yield at 55° C.

In a separate experiment the Molt 4 DNA was 10-fold serially diluted into the cell line GM2064 DNA, containing no β- or δ-globin sequences, available from the Human Genetic Mutant Cell Depository, Camden, N.J., at various concentrations representing varying copies per cell, and amplification was carried out on these samples as described in this example at annealing temperatures of 35° C. and 55° C. At 35° C., the best that can be seen by agarose gel electrophoresis is 1 copy in 50 cells. At 55° C., the best that can be seen is 1/5,000 cells (a 100-fold improvement over the lower temperature), illustrating the importance of increased annealing temperature for Taq polymerase specificity under these conditions.

In a third experiment, DNA from a cell line 368H containing HIV-positive DNA, available from B. Poiesz, State University of New York, Syracuse, N.Y., was similarly diluted into the DNA from the SCl cell line (deposited with ATCC on Mar. 19, 1985; an EBV-transformed β cell line homozygous for the sickle cell allele and lacking any HIV sequences) at various concentrations representing varying copies per cell, and amplification was carried out as described in this Example at annealing temperatures of 35° C. and 55° C., using the primers SK38 and SK39, which amplify a 115 bp region of the HIV sequence:

5'-ATAATCCACCTATCCCAGTAG-GAGAAAT-3'(SK38)

and

5'-TTTGGTCCTTGTCTTATGTCCAGAATGC-3'(SK39)

The results by agarose gel electrophoresis showed that only the undiluted 368H sample could be detected with the annealing temperature at 35° C., whereas at least a 10^2 dilution can be detected with the annealing temperature at 55° C., giving a 100-fold improvement in detection.

The following bacteriophage and bacterial strains were deposited with the Cetus Master Culture Collection, 1400 Fifty-Third Street, Emeryville, Calif., USA (CMCC) and with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Md., USA (ATCC). These deposits were made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for purposes of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture for 30 years from the date of deposit. The organisms will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between applicants and ATCC that assures unrestricted availability upon issuance of the pertinent U.S. patent. Availability of the deposited strains is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

Deposit Designation	CMCC No.	ATCC No.	Deposit
CH35:Taq #4-2	3125	40336	5/29/87
E. coli DG98/pFC83	3128	67422	5/29/87
E. coli DG98/pFC85	3127	67421	5/29/87
E. coli DG95 (λ N_7N_{53} cI_{857} $susP_{80}$)/pFC54.t	2103	39789	8/7/84
E. coli DG116/pAW740CHB	3291	67605	1/12/88

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the cell lines deposited, since the deposited embodiment is intended as a single illustration of one aspect of the invention and any cell lines that are functionally equivalent are within the scope of this invention. The deposit of materials therein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor are the deposits to be construed as limiting the scope of the claims to the specific illustrations that they represent. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

What is claimed is:

1. A recombinant DNA sequence that encodes the thermostable DNA polymerase activity of *Thermus aquaticus.*

2. The recombinant DNA sequence of claim **1** shown in FIG. **1**.

3. A recombinant DNA sequence that encodes amino acid residues 290 through 832 shown in FIG. **1**.

4. The recombinant DNA sequence of claim **3** that encodes amino acid residues 4 through 832 shown in FIG. **1**.

5. A recombinant DNA vector that comprises the DNA sequence of claim **1** and can be used to drive expression of the thermostable DNA polymerase activity of *Thermus aquaticus* in a host cell transformed with the vector.

6. The recombinant vector of claim **5** that is plasmid pLSG1.

7. The recombinant vector of claim **5** that is plasmid pLSG2.

8. The recombinant vector of claim **5** that is plasmid pLSG5.

9. The recombinant vector of claim **5** that is plasmid pLSG6.

10. The recombinant vector of claim **5** that is plasmid pLSG10.

11. A recombinant DNA vector that comprises the DNA sequence of claim **3** and can be used to drive expression of a protein that has thermostable DNA polymerase activity in a host cell transformed with the vector.

12. The recombinant vector of claim **11** that is plasmid pLSG8.

13. The recombinant vector of claim **11** that is plasmid pFC82.

14. The recombinant vector of claim **11** that is plasmid pFC85.

15. A recombinant DNA vector that comprises the DNA sequence of claim **4** and can be used to drive expression of a protein that has thermostable DNA polymerase activity in a host cell transformed with the vector.

16. The recombinant vector of claim **15** that is plasmid pLSG7.

17. A recombinant vector selected from the group consisting of plasmid pFC83, phage CH35:Taq #4-2, and plasmid pSYC1578.

18. A host cell transformed with a vector of claim **5**.

19. The host cell of claim **18** that is *E. coli*/pLSG1.

20. The host cell of claim **18** that is *E. coli*/pLSG2.

21. The host cell of claim **18** that is *E. coli*/pLSG5.

22. The host cell of claim **18** that is *E. coli*/pLSG6.

23. The host cell of claim **18** that is *E. coli*/pLSG7.

24. The host cell of claim **18** that is *E. coli*/pLSG8.

25. The host cell of claim **18** that is *E. coli*/pLSG10.

26. The host cell of claim **18** that is *E. coli*/pFC82.

27. The host cell of claim **18** that is *E. coli*/pFC85.

* * * * *

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 5,079,352
DATED : January 7, 1992
INVENTOR(S) : Gelfand, et al

It is certified that error appears in the above-indentified patent and that said Letters Patent is hereby corrected as shown below:

Column 54,

In line 4, before the period, insert the words --,which polymerase has the amino acid sequence shown in figure 1--.

In lines 5 and 6, cancel Claim 2.

Signed and Sealed this
Twelfth Day of April, 1994

Bruce Lehman

BRUCE LEHMAN
Commissioner of Patents and Trademarks

Attest:

Attesting Officer

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 5,079,352
DATED : January 7, 1992
INVENTOR(S) : Gelfand, et al

It is certified that error appears in the above-indentified patent and that said Letters Patent is hereby corrected as shown below:

In column 54, line 3, delete "activity".

Signed and Sealed this

Second Day of August, 1994

Attest:

BRUCE LEHMAN

Attesting Officer

Commissioner of Patents and Trademarks

United States Patent [19]

Hartley

[11] **Patent Number: 5,035,996**

[45] **Date of Patent: Jul. 30, 1991**

[54] **PROCESS FOR CONTROLLING CONTAMINATION OF NUCLEIC ACID AMPLIFICATION REACTIONS**

[75] Inventor: **James L. Hartley**, Frederick, Md.

[73] Assignee: **Life Technologies, Inc.**, Gaithersburg, Md.

[21] Appl. No.: **360,120**

[22] Filed: **Jun. 1, 1989**

[51] **Int. Cl.**[5] **C12Q 1/68**; C12P 19/34; C12N 9/22; C12N 9/78

[52] **U.S. Cl.** **435/6**; 435/91; 435/200; 435/227

[58] **Field of Search** 435/6, 91, 227, 200; 536/21

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,683,202 7/1987 Mullis 435/91

OTHER PUBLICATIONS

Kwok and Higuchi, "Avoiding False Positives of PCR", *Nature*, vol. 339, (1989) pp. 237–238.

Saiki et al., "Primer-Directed Enzymatic . . . ", *Science*, vol. 239, (1988), 487–491.

Schaaper et al., *PNAS*, "*Infidelity of DNA Synthesis Associated . . . ", vol. 80, pp. 487–491 (1983).*

Primary Examiner—Robert A. Wax
Assistant Examiner—Eric Steffe
Attorney, Agent, or Firm—Sterne, Kessler, Goldstein & Fox

[57] **ABSTRACT**

In the process according to this invention, an amplification procedure is performed on a first sample in which one or more of the four normal ribonucleoside triphosphates (rNTPs) or deoxyribonucleoside triphosphates (dNTPs) is replaced with an exo-sample nucleotide. After amplification, any contaminating amplified product that may be remaining is subjected to a physical, chemical, enzymatic, or biological treatment which renders nucleic acid containing the exo-sample nucleotide substantially unamplifiable. The treatment may be done as a separate step or it may be done in the presence of a second sample containing nucleic acid sequences to be amplified. The amplified nucleic acid sequences derived from the first sample which contaminate the second sample are not further substantially amplified during amplification of nucleic acid sequences of the second sample.

4 Claims, No Drawings

EXHIBIT D

PROCESS FOR CONTROLLING CONTAMINATION OF NUCLEIC ACID AMPLIFICATION REACTIONS

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to processes which amplify nucleic acid sequences. In particular, the present invention discloses a means for eliminating the products of an execution of a nucleic acid amplification process that contaminate subsequent executions of the amplification process.

2. Description of Related Disclosures

The polymerase chain reaction (PCR) procedure amplifies specific nucleic acid sequences through a series of manipulations including denaturation, annealing of primers, and extension of the primers with DNA polymerase (Mullis KB et al., U.S. Pat. No. 4,683,202, U.S. Pat. No. 4,683,195; Mullis KB, EP 201,184; Erlich H, EP 50,424, EP 84,796, EP 258,017, EP 237,362; Erlich H, U.S. Pat. No. 4,582,788; Saiki R et al., U.S. Pat. No. 4,683,202; Mullis KB et al. (1986) in Cold Spring Harbor Symp. Quant. Biol. 51:263; Saiki R et al. (1985) Science 230:1350; Saiki R et al. (1988) Science 231 487; Loh EY et al. (1988) Science 243:217; etc.). (References cited herein are hereby incorporated by reference.) These steps can be repeated many times, potentially resulting in large amplifications of the number of copies of the original specific sequence. It has been shown that even single molecules of DNA can be amplified to produce hundreds of nanograms of product (Li H et al. (1988) Nature 335:414).

Other known nucleic acid amplification procedures include the transcription-based amplification system of Kwoh D et al. (1989) Proc. Natl. Acad. Sci. USA 86:1173.

"Uracil DNA glycosylase" (UDG), a term of art, refers to an enzyme which cleaves the glycosidic bond between the base uracil and the sugar deoxyribose, only when the monomeric nucleotide dUTP is incorporated into a DNA molecule, resulting in incorporation of a deoxyuridine moiety (Duncan B (1981) in *The Enzymes* 14:565, ed.: Boyer P). The enzyme does not act upon free dUTP, free deoxyuridine, or RNA (Duncan, supra).

A consequence of amplification processes such as PCR is that the amplification products themselves can be substrates for subsequent PCR procedures. Furthermore, because the quantities of the amplification products can be large, the dispersal of even an extremely small fraction of a reaction such as a PCR reaction into the laboratory area potentially can lead to contamination of later attempts to amplify other samples.

The present invention represents an improvement upon in vitro nucleic acid amplification procedures in general by making amplification products distinguishable from naturally occurring DNA. Accordingly, such products are rendered inactive as templates for further amplification prior to the start of the succeeding amplification reaction.

SUMMARY OF THE INVENTION

The present invention involves a process for incorporating an exo-sample nucleotide into DNA or RNA during amplification procedures. The invention eliminates the products of previous amplifications from further amplification by means of a treatment that leaves

nucleic acid from the sample unaffected in its ability to be amplified. This treatment greatly reduces a major problem associated with amplification of nucleic acids, namely contamination of starting materials with the end products of previous amplification processes. In other words, the present invention provides a process of discriminating against amplification products, and in favor of nucleic acids normally found in nature, prior to the start of succeeding amplification reactions.

More specifically, the invention relates to in vitro procedures which utilize enzymes to amplify specific nucleic acid sequences. One example of such a procedure is known as the polymerase chain reaction (PCR). A serious limitation of the PCR procedure and other similar procedures is contamination of the laboratory environment with the amplified nucleic acid end products of individual reactions. Such contamination commonly results in amplification not only of authentic nucleic acid which may be present in the sample of interest, but also of the contaminating end products from previous reactions. The present invention provides a process to remove possible contamination of this type, without affecting the desired amplification of authentic nucleic acids.

The present invention involves first performing amplification procedures in which one or more of the four normal ribonucleoside triphosphates (rNTPs) or deoxyribonucleoside triphosphates (dNTPs) is replaced with one or more exo-sample nucleotides that are normally absent from or present very rarely in nucleic acids found in the samples whose amplification is desired. The DNA or RNA produced during such amplification processes can be differentiated from sample nucleic acids. Thus, one can discriminate against nucleic acids produced during amplification processes in favor of sample DNA or RNA prior to or during succeeding amplification processes, such that previously amplified nucleic acid can no longer be amplified, while sample DNA or RNA remains amplifiable.

It is believed that since the invention of the various nucleic acid amplification methods no one has disclosed a means for eliminating contamination of input nucleic acid by the products of previous cycles of amplification.

DETAILED DESCRIPTION OF THE INVENTION

The term "amplifying", as used herein, refers to any in vitro process for increasing the number of copies of a nucleotide sequence or sequences. Nucleic acid amplification results in the incorporation of nucleotides into DNA or RNA.

"Nucleotide" is a term of art that refers to a base-sugarphosphate combination. Nucleotides are the monomeric units of nucleic acid polymers, i.e. of DNA and RNA. The term includes ribonucleoside triphosphates, such as rATP, rCTP, rGTP, or rUTP, and deoxyribonucleoside triphosphates, such as dATP, dCTP, dGTP, or dTTP.

"Nucleoside" is a term of art referring to a base-sugar combination, i.e. a nucleotide lacking a phosphate moiety. It is recognized in the art that there is a certain inter-changabilty in usage of the terms nucleoside and nucleotide. For example, the nucleotide deoxyuridine triphosphate, dUTP, is a deoxyribonucleoside triphosphate. After incorporation into DNA, it serves as a DNA monomer, formally being deoxyuridylate, i.e. dUMP or deoxyuridine monophosphate. One may say

that one incorporates dUTP into DNA even though there is no dUTP moiety in the resultant DNA. Similarly, one may say that one incorporated deoxyuridine into DNA even though that is only a part of the substrate molecule.

The term "exo-sample nucleotide", as used herein, refers to a nucleotide which is generally not found in the sample to be amplified. For most DNA samples, deoxyuridine is an example of an exo-sample nucleotide. Although the triphosphate form of deoxyuridine, dUTP, is present in living organisms as a metabolic intermediate, it is rarely incorporated into DNA. When dUTP is accidentally incorporated into DNA, the resulting deoxyuridine is promptly removed in vivo by normal processes, e.g. processes involving the enzyme UDG. Thus, deoxyuridine occurs rarely or never in natural DNA. It is recognized that some organisms may naturally incorporate deoxyuridine into DNA. For nucleic acid samples of those organisms, deoxyuridine would not be considered an exo-sample nucleotide. The presence of deoxyuridine, or any other exo-sample nucleotide, may be determined readily using methods well known to the art. Other exo-sample nucleotides may be envisioned. Numerous DNA glycosylases are known to the art. An exo-sample nucleotide which may be incorporated into DNA during an amplification and a DNA glycosylase that acts on it may be used in this invention. Similarly, bromodeoxyuridine (BdUR) is well known in the art to be incorporated into DNA. DNA containing BdUR may be degraded on exposure to light under appropriate conditions.

The term "incorporating" refers to becoming part of a nucleic acid polymer.

The term "terminating" refers herein to causing a treatment to stop. The term includes means for both permanent and conditional stoppages. For example, if the treatment is enzymatic, both permanent heat denaturation and lack of enzymatic activity due to a temperature outside the enzyme s active range would fall within the scope of this term.

In the process according to this invention, an amplification procedure is performed on a first sample in which one or more of the four normal ribonucleoside triphosphates (rNTPs) or deoxyribonucleoside triphosphates (dNTPs) is replaced with an exo-sample nucleotide. After amplification, any contaminating amplified product that may be remaining is subjected to a physical, chemical, enzymatic, or biological treatment which renders nucleic acid containing the exo-sample nucleotide substantially unamplifiable. The treatment may be done as a separate step, or preferably, may be done in the presence of a second sample containing nucleic acid sequences to be amplified. The amplified nucleic acid sequences derived from the first sample which contaminate the second sample are not further substantially amplified during amplification of nucleic acid sequences of the second sample.

The deoxyribonucleoside triphosphate dUTP exemplifies an exo-sample nucleotide which may be conveniently incorporated into an enzymatic DNA amplification procedure, exemplified herein by PCR, thereby resulting in deoxyuridine-containing DNA. The DNA products of such a reaction will normally contain many uracil bases. Discrimination between natural DNA and the resultant, deoxyuridine-containing products of amplification procedures may be obtained with the enzyme uracil DNA glycosylase (UDG). Treatment of DNA containing uracil bases with uracil DNA glycosylase results in cleavage of the glycosidic bond between the deoxyribose of the DNA sugar-phosphate backbone and the uracil base. The loss of the uracil creates an apyrimidinic site in the DNA, which blocks DNA polymerase from using the DNA strand as a template for the synthesis of a complementary DNA strand (Schaaper R et al. (1983) Proc. Natl. Acad. Sci. USA 80:487). The presence of substantial numbers of apyrimidinic sites in each DNA target molecule interferes with amplification procedures which use DNA polymerase to synthesize copies of target DNA.

As exemplified herein, the basic amplification protocol is the well known PCR method. PCR was modified in three ways: (1) dUTP was substituted for dTTP; (2) UDG was added to the initial PCR reaction mixture; and (3) an initial incubation period was added to allow UDG to destroy contaminating products of prior PCR reactions. The UDG itself was either permanently inactivated by high temperature in the first PCR cycle or was not active at the high temperatures used with Taq polymerase in the currently preferred PCR protocol. This inactivation prevents UDG from destroying newly-synthesized PCR products. Nucleic acid amplification protocols that do not eliminate UDG activity usually will require an extra UDG-inactivation step.

While termination of the physical, chemical, enzymatic, or biological treatment that renders nucleic acid containing the exo-sample nucleotide resistant to the amplification process is preferred (as exemplified herein, heat inactivation of UDG), the invention also includes embodiments lacking a termination step. For example, one might use amounts of enzyme and durations of treatment high enough to eliminate expected contamination of starting materials but insufficient to keep up with the rate of amplification. In other words, a treatment might be able to destroy contaminating nucleic acid but an amplification process might still be able to produce new nucleic acid faster than the treatment could destroy the newly synthesized nucleic acid.

Variations on the herein disclosed invention may also be envisioned. For example, the amplification may be done without an exo-sample nucleotide, i.e., using normal nucleotides. A normal nucleotide in the amplified DNA is then converted into an exo-sample nucleotide. The converted DNA can then be removed from any samples which it later contaminates. An example would be the conversion of neighboring pyrimidine residues, especially thymidine, into pyrimidine dimers (thymidine dimers), which make DNA unsuitable as a template. Thymidine dimers can also be removed by enzymes such as exonuclease VII and recBC.

EXAMPLE

A polymerase chain reaction (PCR) was performed to amplify a region of the human papilloma virus type 16 (HPV 16) DNA (Durst M et al. (1983) Proc. Natl. Acad. Sci. USA 80:3812). The sequences of the primers used were 5'GGTCGATGTATGTCTTGTTG3' and 5'GTCTACGTGTGTGCTTTGTAC3'.

HPV 16 DNA was excised from a full length plasmid clone, pT7HPV16 (for the purposes of this invention, equivalent to the pUC8 plasmids described by Seedoff K et al. (1985) Virol. 145:181) with the restriction enzyme BamH 1. The linear DNA (10 picograms) was added to PCR reactions containing 50 microliters of 25 mM Tris HCl pH 8.3, 5 mM $MgCl_2$, 50 mM NaCl, 0.01% gelatin, 0.05% W1 polyoxyethylene ether detergent (Sigma), 0.2 mM each dATP, dGTP, dCTP, 0.2

mM either dUTP or dTTP, 1 micromolar of each primer, and 12.5 units of thermostable DNA polymerase from *Thermus aquaticus* (Cetus/Perkin-Elmer). The reactions were amplified in a thermal cylcer (Cetus/Perkin-Elmer) using the following temperature profile: 5 minutes at 94° C., then 30 cycles of 1 minute at 94° C. (denaturation), two minutes at 55° C. (annealing), and 3 minutes at 72° C. (primer extension). After completion of the temperature cycles, a final extension of 10 minutes at 72° C. was done. Amplification of the 284 base pair HPV 16 DNA fragment was confirmed by agarose/ethidium bromide gel electrophoresis (Maniatis T et al. (1982) *Molecular Cloning,* Cold Spring Harbor Laboratory) of the PCR reaction products (5 microliters of each reaction per lane). All reactions showed substantial amplification. Negative control reactions to which no HPV 16 DNA was added did not produce any DNA products.

The concentration of the PCR amplification products was estimated from the agarose gel. New PCR reactions were contaminated with ten femtogram quantities of the amplification products that contained either deoxythymidine, resulting from incorporation of dTTP, or deoxyuridine, from dUTP-containing reactions. Positive control reactions contained 10 picograms of linear HPV 16 DNA. Negative control reactions did not receive any target DNA. The new PCR reactions contained dUTP instead of dTTP, and either 5 nanograms of UDG (Van de Sande H, University of Calgary; also available from Duncan Laboratories, 19 E. Central Ave., Paoli, PA 19301 USA) or no UDG. All reactions were incubated for 15 minutes at 37° C. to allow the UDG to act on deoxyuridine-containing DNA, and then were taken through the same thermal cycling protocol as above. Aliquots of each reaction were analyzed by agarose/ethidium bromide gel electrophoresis.

The agarose gel analysis showed that without UDG treatment the deoxyuridine-containing PCR products could be re-amplified to give a DNA product indistinguishable in size, as evidenced by gel electrophoresis, from the products obtained by amplifying the normal HPV 16 DNA. Reactions in which the deoxyuridine-containing DNA was incubated with UDG prior to PCR did not give any visible products on the agarose gel. PCR amplification products that contained deoxythymidine were successfully amplified whether or not they had been incubated with UDG. This experiment showed that UDG substantially abolished amplification of PCR products containing deoxyuridine, but had no substantial effect on the amplification of DNA containing deoxythymidine.

Although the foregoing refers to particular preferred embodiments, it will be understood that the present invention is not so limited. It will occur to those of ordinary skill in the art that various modifications may be made to the disclosed embodiments and that such modifications are intended to be within the scope of the present invention.

What is claimed is:

1. A method for controlling contamination in sequential nucleic acid amplification processes comprising first and second nucleic acid amplification processes to amplify nucleic acid sequences in a first and a second sample, respectively, which comprises:

carrying out said first nucleic acid amplification process on said nucleic acid sequence in said first sample in the presence of deoxyuridine triphosphate, to thereby generate a deoxyuracil-containing product of amplification; and

treating said deoxyuracil-containing product of amplification with uracil DNA glycosylase prior to carrying out said second amplification process on said nucleic acid sequence in said second sample.

2. The process of claim **1** wherein said nucleic acid sequences in said first and second samples have two separate complementary strands each,

wherein said first nucleic acid amplification process on said first sample comprises:

(1) treating said strands with two oligonucleotide primers, for each different specific sequence being amplified, under conditions such that for each different sequence being amplified, an extension product of each primer is synthesized which is complementary to each nucleic acid strand and which incorporates deoxyuridine, wherein said primers are selected so as to be sufficiently complementary to different strands of each specific sequence to hybridize therewith, such that the extension product synthesized from one primer, when it is separated from its complement, can serve as a template for synthesis of the extension product of the other primer;

(2) separating the primer extension products from the templates on which they are synthesized to produce single-stranded molecules;

(3) treating said single-stranded molecules generated from step (2) with the primers of step (1) under conditions that a primer extension product is synthesized using each of the single strands produced in step (2) as a template;

(4) repeating steps (1) through (3) at least once, thereby amplifying said specific nucleic acid sequence contained in said first sample;

and wherein said second nucleic acid amplification process on said second sample comprises

(5) treating said second sample with uracil DNA glycosylase; and

(6) repeating steps (1) through (3) on said second sample at least once, thereby amplifying the nucleic acid sequence contained in said second sample.

3. The process of claim **1** wherein, after treating said deoxyuracil-containing product with uracil DNA glycosylase, and prior to carrying out said second amplification process, the second sample is heated so as to terminate the action of said uracil DNA glycosylase.

4. The process of claim **2** wherein, after said step (5), the process comprises:

(7) heating said sample so as to terminate the action of said uracil DNA glycosylase.

* * * * *

US005945313A

United States Patent [19]

Hartley et al.

[11] **Patent Number: 5,945,313**

[45] **Date of Patent: *Aug. 31, 1999**

[54] **PROCESS FOR CONTROLLING CONTAMINATION OF NUCLEIC ACID AMPLIFICATION REACTIONS**

[75] Inventors: **James L. Hartley**, Frederick; **Mark Berninger**, Gaithersburg, both of Md.

[73] Assignee: **Life Technologies, Inc.**, Rockville, Md.

[*] Notice: This patent is subject to a terminal disclaimer.

[21] Appl. No.: **08/962,701**

[22] Filed: **Nov. 3, 1997**

Related U.S. Application Data

[63] Continuation of application No. 08/221,465, Apr. 1, 1994, Pat. No. 5,683,896, which is a continuation of application No. 08/079,835, Jun. 22, 1993, abandoned, which is a continuation of application No. 07/728,874, Jul. 12, 1991, abandoned, which is a continuation-in-part of application No. 07/633,389, Dec. 31, 1990, abandoned, which is a continuation-in-part of application No. 07/360,120, Jun. 1, 1989, Pat. No. 5,035,996, and application No. 07/401,840, Sep. 1, 1989, abandoned.

[51] **Int. Cl.**[6] **C12P 19/34**; C12Q 1/68; C12N 9/14; C12N 9/12

[52] **U.S. Cl.** **435/91.2**; 435/6; 435/195; 435/810; 435/194; 935/77; 935/78

[58] **Field of Search** 435/6, 91.2, 195, 435/810, 194; 935/72, 78

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,683,195	7/1987	Mullis et al.	435/6
4,683,202	7/1987	Mullis	435/91
4,873,192	10/1989	Kunkel	435/172.3
4,876,187	10/1989	Duck et al.	536/27
4,965,188	10/1990	Mullis et al.	435/6
5,035,996	7/1991	Hartley	435/6
5,137,814	8/1992	Rashtchian et al.	435/91
5,334,515	8/1994	Rashtchian et al.	435/91.2
5,418,149	5/1995	Gelfand et al.	435/91.2
5,427,929	6/1995	Richards et al.	435/91.2

FOREIGN PATENT DOCUMENTS

59577/90	1/1991	Australia .
0 123 513	10/1984	European Pat. Off. .
0 200 362	12/1986	European Pat. Off. .
0 237 362	9/1987	European Pat. Off. .
0 258 017	3/1988	European Pat. Off. .
0 320 308	6/1989	European Pat. Off. .
0 336 731	10/1989	European Pat. Off. .
0 401 037	12/1990	European Pat. Off. .
0 407 291	1/1991	European Pat. Off. .
0 415 755	3/1991	European Pat. Off. .
WO 88/10315	12/1988	WIPO .
WO 89/09835	10/1989	WIPO .
WO 89/11548	11/1989	WIPO .
WO 91/17270	11/1991	WIPO .

OTHER PUBLICATIONS

Barany, F., "Genetic disease detection and DNA amplification using cloned thermostable ligase," *Proc. Natl. Acad. Sci. USA 88*:189–193 (1991).

Cimino, G.D. et al., "More false–positive problems," *Nature 345*:773–774 (1990).

Delort, A.–M. et al., "Excision of uracil residues in DNA: mechanism of action of *Escherichia coli* and *Micrococcus luteus* uracil–DNA glycosylases," *Nucleic Acids Res. 13*(2):319–335 (1985).

Doetsch, P.W., and Cunningham, R.P., "The enzymology of apurinic/apyrimidinic endonucleases," *Mutation Res. 236*:173–201 (1990).

Duncan, B.K., "DNA Glycosylases," *The Enzymes XIV*:565–586 (1981).

Kwoh, D.Y. et al., "Transcription–based amplification system and detection of amplified human immunodeficiency virus type 1 with a bead–based sandwich hybridization format," *Proc. Natl. Acad. Sci USA 86*:1173–1177 (Feb. 1989).

Kwok, S. and Higuchi, R., "Avoiding false positives with PCR," *Nature 339*:237–238 (May 1989).

Li, H. et al., "Amplification and analysis of DNA sequences in single human sperm and diploid cells," *Nature 335:*414–417 (Sep. 1988).

Loh, E.Y. et al., "Polymerase Chain Reaction with Single–Sided Specificity: Analysis of T Cell Receptor σ chain," *Science 243:*217–220 (Jan. 1989).

Longo, M.C. et al., "Use of uracil DNA glycosylase to control carry–over contamination in polymerase chain reactions," *Gene 93:*125–128 (1990).

Mullis, K. et al., "Specific Enzymatic Amplification of DNA In Vitro: The Polymerase Chain Reaction," *Cold Spring Harb. Symp. Quant. Biol. LI*:263–273 (1986).

Perkin–Elmer Cetus Brochure, "GeneAmp® PCR Carry–over Prevention Kit," B–3, p. 2 (1990).

Saiki, R.K. et al., "Enzymatic Amplification of β–Globin Genomic Sequences and Restriction Site Analysis for Diagnosis of Sickle Cell Anemia," *Science 230:*1350–1354 (1985).

Saiki, R.K. et al., "Primer–Directed Enyzmatic Amplification of DNA with a Thermostable DNA Polymerase," *Science 239*:487–491 (Jan. 1988).

(List continued on next page.)

Primary Examiner—Stephanie Zitomer
Attorney, Agent, or Firm—Sterne, Kessler, Goldstein & Fox P.L.L.C.

[57] **ABSTRACT**

This invention relates to a method of incorporating an exo-sample nucleotide into the amplified product strands resulting from a nucleic acid amplification process. Once the product strands have been obtained and analyzed (e.g., by hybridization, Southern blot, etc.), the exo-sample strands can be selectively destroyed by acting on the incorporated exo-sample nucleotide. Two embodiments are presented. In a first embodiment, the exo-sample nucleotide is incorporated by carrying out the amplification reaction in the presence of an excess of exo-sample nucleotide triphosphate. In a second embodiment, the exo-sample nucleotide is incorporated by carrying out the amplification reaction in the presence of an oligonucleotide which has, as part of its sequence, one or more exo-sample nucleotides.

61 Claims, 1 Drawing Sheet

OTHER PUBLICATIONS

Schaaper, R.M. et al., "Infidelity of DNA synthesis associated with bypass of apurinic sites," *Proc. Natl. Acad. Sci. USA 80*: 487–491 (1983).

Varshney, U. and van de Sande, J.H., "Specificities and Kinetics of Uracil Excision from Uracil–Containing DNA Oligomers by *Escherichia coli* Uracil DNA Glycosylase," *Biochemistry 30*:4055–4061 (1991).

Wu, D.Y. and Wallace, R.B., "The Ligation Amplification Reaction (LAR)–Amplification of Specific DNA Sequences Using Sequential Rounds of Template–Dependent Ligation," *Genomics 4:*560–569 (May 1989).

European Search Report for Appl. No. EP 90 30 9492 corresponding to U.S. Appl. No. 07/401,840.

European Search Report for Appl. No. EP 90 30 6001 corresponding to U.S. Appl. No. 07/360,120.

Almoguera, C. et al., "Most Human Carcinomas of the Exocrine Pancreas Contain Mutant c–K–ras Genes," *Cell 53*(*4*):549–554 (May 1988).

Bos, J.L. et al., "Prevalence of ras gene mutations in human colorectal cancers," *Nature 327*:293–297 (1987).

Engelke, D.R. et al., "Direct sequencing of enzymatically amplified human genomic DNA," *Proc. Natl. Acad. Sci. USA 85*(*2*):544–548 (Jan. 1988).

Erlich, H.A. et al., "Specific DNA amplification," *Nature 331:*461–462 (Feb. 1988).

Farr, C.J. et al., "Analysis of RAS gene mutations in acute myeloid leukemia by polymerase chain reaction and oligonucleotide probes," *Proc. Natl. Acad. Sci. USA 85*(*5*):1629–1633 (Mar. 1988).

Higuchi, R. et al., "DNA typing from single hairs," *Nature 332*:543–546 (Apr. 1988).

Higuchi, R. et al., "A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions," *Nucleic Acids Res. 16*(*15*):7351–7367 (Aug. 1988).

Kornberg, A. and Baker, T.A., *DNA Replication,* W.H. Freeman, pub., NY, pp. 136–137 (1992).

Kwok, S. et al., "Identification of Human Immunodeficiency Virus Sequences by Using In Vitro Enzymatic Amplication and Oligomer Cleavage Detection," *J. Virol 61*(*5*):1690–1694 (1987).

Langer, P.R. et al., "Enzymatic synthesis of biotin–labeled polynucleotides: Novel nucleic acid affinity probes," *Proc. Natl. Acad. Sci. USA 78*(*11*):6633–6637 (1981).

Lindahl, T. et al., "DNA N–Glycosidases. Properties of Uracil–DNA Glycosidase from *Escherichia coli," J. Biol. Chem. 252*(*10*):3286–3294 (1977).

Marx, J.L., "Multiplying Genes by Leaps and Bounds," *Science 240:*1408–1410 (Jun. 1988).

Maxam, A.M. and Gilbert, W., "A new method for sequencing DNA," *Proc. Natl. Acad. Sci. USA 74*(*2*):560–564 (1977).

Mole, S.E. et al., "Using the polymerase chain reaction to modify expression plasmids for epitope mapping," *Nucleic Acids Res. 17*(*8*):3319 (Apr. 1989).

Mullis, K.B. and Faloona, F.A., "Specific Synthesis of DNA in Vitro via a Polymerase–Catalyzed Chain Reaction," *Methods Enymol. 155*:355–350 (1987).

Orlandi, R. et al., "Cloning immunoglobulin variable domains for expression by the polymerase chain reaction," *Proc. Natl. Acad. Sci. USA 86*(*10*):3833–3837 (May 1989).

Oste, C., "Polymerase Chain Reaction," *BioTechniques 6*(*2*):162–167 (Feb. 1988).

Ou, C.–Y. et al., "DNA Amplification for Direct Detection of HIV–1 in DNA of Peripheral Blood Mononuclear Cells," *Science 239*:295–297 (Jan. 1988).

Syvänen, A.–C. et al., "Quantification of polymerase chain reaction products by affinity–based hybrid collection," *Nucleic Acids Res. 16*(*23*):11327–11338 (Dec. 1988).

Vosberg, H.–P., "The polymerase chain reaction: an improved method for the analysis of nucleic acids," *Hum. Genet. 83*(*1*):1–15 (Aug. 1989).

Wong, C. et al., "Characterization of β–thalassaemia mutations using direct genomic sequencing of amplified single copy DNA," *Nature 330:*384–386 (1987).

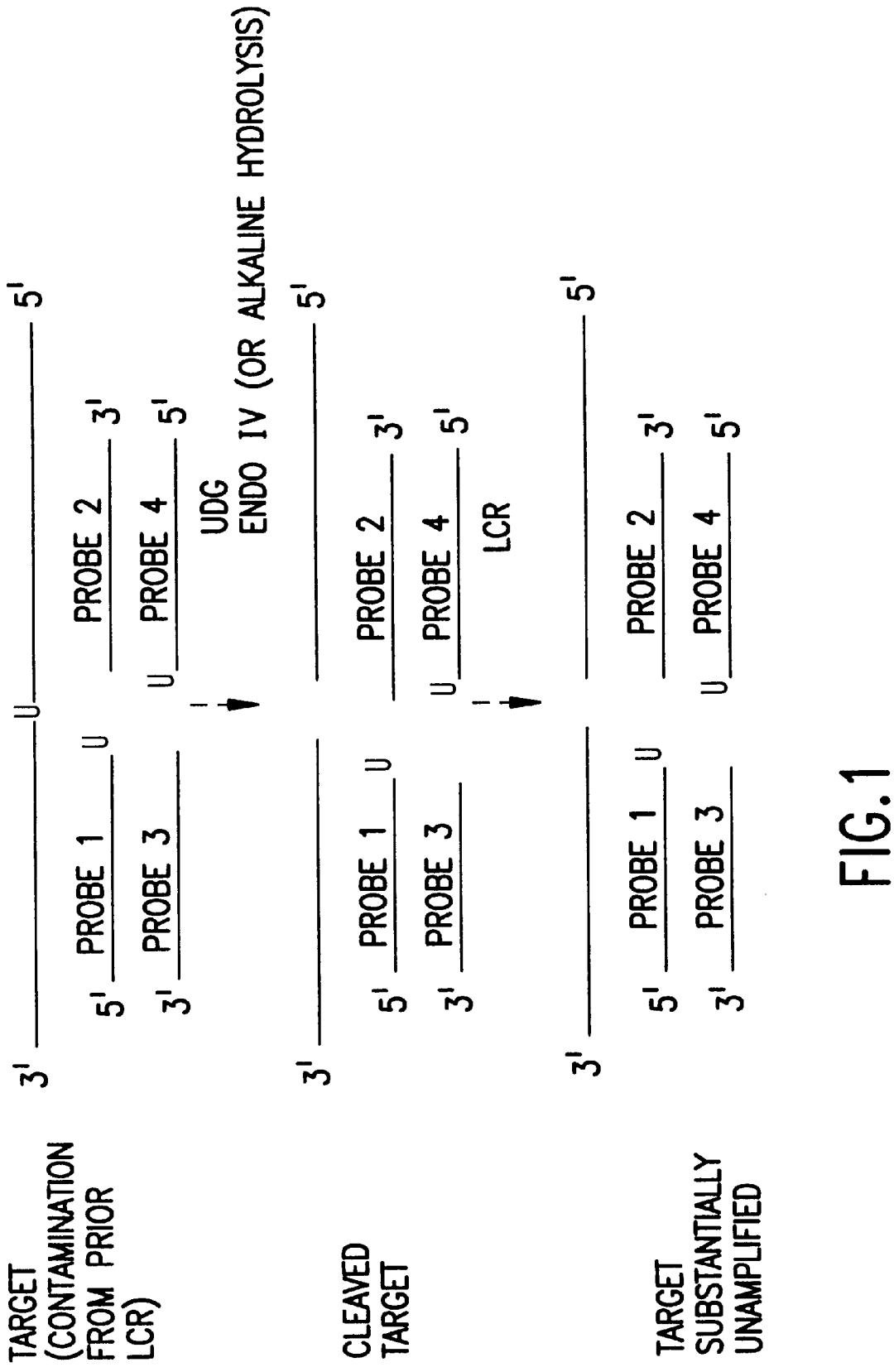

TARGET (CONTAMINATION FROM PRIOR LCR)
CLEAVED TARGET
TARGET SUBSTANTIALLY UNAMPLIFIED
PROBE 1
PROBE 2
PROBE 3
PROBE 4
UDG
ENDO IV (OR ALKALINE HYDROLYSIS)
LCR
5'
3'
U

FIG. 1

PROCESS FOR CONTROLLING CONTAMINATION OF NUCLEIC ACID AMPLIFICATION REACTIONS

This application is a continuation of U.S. application Ser. No. 08/221,465 filed Apr. 1, 1994 (now U.S. Pat. No. 5,683,896), which is a continuation of U.S. application Ser. No. 08/079,835 filed Jun. 22, 1993 (now abandoned), which is a continuation of U.S. application Ser. No. 07/728,874 filed Jul. 12, 1991 (now abandoned), which is a continuation-in-part of U.S. application Ser. No. 07/633,389 filed Dec. 31, 1990 (now abandoned), which is a continuation-in-part of U.S. application Ser. No. 07/360,120 filed on Jun. 1, 1989 (now U.S. Pat. No. 5,035,996), and Ser. No. 07/401,840 filed on Sep. 1, 1989 (now abandoned).

FIELD OF THE INVENTION

The present invention relates to an improvement in processes which amplify nucleic acid sequences. In particular, invention is directed to a means for eliminating the products of an execution of a nucleic acid amplification process that contaminate subsequent executions of the amplification process.

One embodiment is directed to a method for preventing carryover-contamination of nucleic acid samples by eliminating the products of an execution of an oligonucleotide-dependent nucleic acid amplification process that contaminate subsequent executions of the amplification process. Such oligonucleotide-dependent nucleic acid amplification processes may include, for example, Polymerase Chain Reaction (PCR) and Ligase Chain Reaction (LCR).

The improvements of the present invention ensure that the results of an amplification process do not reflect the presence of carryover-contaminating nucleic acid template.

BACKGROUND OF THE INVENTION

The polymerase chain reaction (PCR) procedure amplifies specific nucleic acid sequences through a series of manipulations including denaturation, annealing of oligonucleotide primers, and extension of the primers with DNA polymerase (Mullis, K. B. et al., U.S. Pat. No. 4,683,202, U.S. Pat. No. 4,683,195; Mullis, K. B., EP 201,184; Erlich, H., EP 50,424, EP 84,796, EP 258,017, EP 237,362; Erlich, H., U.S. Pat. No. 4,582,788; Saiki, R. et al., U.S. Pat. No. 4,683,202; Mullis, K. B. et al. Cold Spring Harbor *Symp. Quant. Biol.* 51:263 (1986); Saiki, R. et al. *Science* 230:1350 (1985); Saiki, R. et al. *Science* 231:487 (1988); Loh, E. Y. et al. *Science* 243:217 (1988)). These steps can be repeated many times, potentially resulting in large amplification of the number of copies of the original specific sequence. It has been shown that even a single copy of a DNA sequence can be amplified to produce hundreds of nanograms of product (Li, H. et al *Nature* 335:414 (1988)).

Other known nucleic acid amplification procedures include transcription-based amplification systems (Kwoh, D. et al. *Proc. Natl. Acad. Sci. USA* 86:1173 (1989); Gingeras, T. R. et al., WO 88/10315).

Schemes based on ligation of two (or more) oligonucleotides in the presence of a nucleic acid target having the sequence of the resulting "di-oligonucleotide," thereby amplifying the di-oligonucleotide, are also known (Wu, D. Y. and Wallace, R. B. *Genomics* 4:560 (1989); Backman et al., EP 320,308; Wallace, B., EP 336,731; and Orgel, L., WO 89/09835). Such oligonucleotide-dependent amplifications are termed "Ligase Chain Reaction" (LCR).

A consequence of amplification processes, such as PCR or LCR, is that the amplification products themselves can be

substrates for subsequent PCR or LCR procedures. Furthermore, because the quantities of the amplification products can be large, and because the sensitivity of PCR and LCR is so great, the dispersal of even an extremely small fraction of a reaction, such as a PCR or LCR reaction, into the laboratory area potentially can lead to contamination of later attempts to amplify other samples, thereby resulting in false positives. Extreme care must be taken to avoid carryover contamination (Kwok, S. and Higuchi, R. *Nature* 339:237 (1989)); this is very inconvenient and adds significantly to the cost of doing amplifications such as PCR and LCR.

Thus a need exists for a routine, economical method of nucleic acid amplification wherein such amplification may be performed without concern as to possible carryover-contamination from previous amplifications.

The invention represents an improvement upon in vitro nucleic acid amplification procedures in general by making amplification products distinguishable from naturally occurring DNA. Accordingly, such products are rendered inactive as templates for further amplification prior to the start of the succeeding amplification reaction.

SUMMARY OF THE INVENTION

This invention relates to a method of incorporating an exo-sample nucleotide into the amplified product strands resulting from a nucleic acid amplification process. Once the product strands have been obtained and analyzed (e.g., by hybridization, Southern blot, etc.), the exo-sample strands can be selectively destroyed by acting on the incorporated exo-sample nucleotide.

Two embodiments are presented. In a first embodiment, the exo-sample nucleotide is incorporated by carrying out the amplification reaction in the presence of an excess of exo-sample nucleotide triphosphate.

In a second embodiment, the exo-sample nucleotide is incorporated by carrying out the amplification reaction in the presence of an oligonucleotide which has, as part of its sequence, one or more exo-sample nucleotides. The primer containing exo-sample nucleotide(s) can be used alone or in combination with the first embodiment, i.e., also incorporating the exo-sample nucleotide by carrying out the amplification reaction in the presence of an excess of exo-sample nucleotide triphosphate.

In a variation of the second embodiment, the exo-sample nucleotide is incorporated in at least one oligonucleotide before the amplification reaction. Preferably the exo-sample nucleotide is incorporated at or near the oligonucleotide termini. Before amplification, the oligonucleotide containing exo-sample nucleotide is substantially amplifiable. After amplification, the amplified oligonucleotide (containing exo-sample nucleotide) is substantially unamplifiable. The oligonucleotide containing exo-sample nucleotide may be made unamplifiable by a treatment during the amplification process. Causing the amplified oligonucleotide containing exo-sample nucleotide to be cleaved at or near the location of the exo-sample nucleotide is one example of making such an oligonucleotide substantially unamplifiable.

The invention eliminates the products of previous amplifications from further amplification by means of a treatment that leaves nucleic acid from the sample unaffected in its ability to be amplified. This treatment greatly reduces a major problem associated with amplification of nucleic acids, namely contamination of starting materials with the end products of previous amplification processes. In other words, this invention provides a process of discriminating

against amplification products, and in favor of nucleic acids normally found in nature, prior to the start of succeeding amplification reactions.

More specifically, this invention relates to in vitro procedures which utilize enzymes to amplify specific nucleic acid sequences. Examples of such procedures include polymerase chain reaction (PCR) and ligase chain reaction (LCR). A serious limitation of the PCR procedure, the LCR procedure and other similar procedures is contamination of the laboratory environment with the amplified nucleic acid end products of individual reactions. Such contamination commonly results in amplification not only of authentic nucleic acid which may be present in the sample of interest, but also of the contaminating end products from previous reactions. This invention provides a process to remove possible contamination of this type, without affecting the desired amplification of authentic nucleic acids.

The first embodiment involves first performing amplification procedures in which one or more of the four normal ribonucleoside triphosphates (rNTPs) or deoxyribonucleoside triphosphates (dNTPs) is replaced with one or more exo-sample nucleotides that are normally absent from or present very rarely in nucleic acids found in the samples whose amplification is desired. The DNA or RNA produced during such amplification processes can be differentiated from sample nucleic acids. Thus, one can discriminate against nucleic acids produced during amplification processes in favor of sample DNA or RNA prior to or during succeeding amplification processes, such that previously amplified nucleic acid can no longer be amplified, while sample DNA or RNA remains amplifiable.

The present invention represents an improvement upon in vitro oligonucleotide-dependent, nucleic acid amplification procedures. In the methods of the present invention, amplification products are made distinguishable from the nucleic acid substrate used to initiate the amplification in a manner which imparts distinct properties to the amplification products. Accordingly, prior to the start of a new amplification reaction, these distinct properties may be exploited so as to render former amplification products inactive as templates in subsequent amplification reactions.

Therefore, the present invention is directed to:

a process for oligonucleotide-dependent amplification of one or more nucleic acid sequences in a sample, comprising the steps of:

(a) amplifying nucleic acid of a first sample, wherein said amplifying is dependent on one or more specific oligonucleotides, and wherein at least one of such specific oligonucleotides comprises an exo-sample nucleotide, thereby producing amplified nucleic acid containing an exo-sample nucleotide, and

(b) subjecting the nucleic acid of a second sample to treatment which renders said amplified nucleic acid containing an exo-sample nucleotide substantially unamplifiable in an amplification dependent on such specific oligonucleotides and which does not substantially affect amplification of nucleic acid that does not contain the exo-sample nucleotide;

whereby amplified nucleic acid sequences derived from such first sample which contaminate such second sample are not further substantially amplified during amplification dependent on such specific oligonucleotides of the nucleic acid sequences of such second sample.

The present invention is further directed to a process for oligonucleotide-dependent amplification of one or more nucleic acid sequences in a sample, comprising the steps of:

(a) amplifying nucleic acid of a first sample, wherein such amplifying is dependent on one or more specific oligonucleotides, and wherein at least one of said specific oligonucleotides comprises an exo-sample nucleotide, thereby producing amplified nucleic acid containing an exo-sample nucleotide, and

(b) subjecting the nucleic acid of a second sample to treatment which renders said amplified nucleic acid containing an exo-sample nucleotide substantially unamplifiable in an amplification dependent on said specific oligonucleotides and which does not substantially prevent amplification of nucleic acid that does not contain said exo-sample nucleotide;

wherein said oligonucleotide comprising an exo-sample nucleotide before such amplification is substantially less susceptible to said treatment than said amplified nucleic acid containing an exo-sample nucleotide, whereby amplified nucleic acid sequences derived from said first sample which contaminate said second sample are not further substantially amplified during amplification dependent on said specific oligonucleotides of the nucleic acid sequences of said second sample.

The present invention is further directed to a process for amplifying at least one specific nucleic acid sequence contained in a nucleic acid or a mixture of nucleic acids wherein each nucleic acid consists of two separate complementary strands, of equal or unequal length, which process further comprises:

(e) treating the strands with two oligonucleotide primers, for each different specific sequence being amplified, under conditions such that for each different sequence being amplified an extension product of each primer is synthesized which is complementary to each nucleic acid strand, wherein such primers are selected so as to be sufficiently complementary to different strands of each specific sequence to hybridize therewith such that the extension product synthesized from one primer, when it is separated from its complement, can serve as a template for synthesis of the extension product of the other primer;

(f) separating the primer extension products from the templates on which they were synthesized to produce single-stranded molecules; and

(g) treating the single-stranded molecules generated from step (f) with the primers of step (e) under conditions that a primer extension product is synthesized using each of the single strands produced in step (f) as a template.

Exo-sample nucleotides can, according to the invention, be incorporated during amplification steps (e) and (g) or the exo-sample nucleotide can be incorporated in at least one primer used in step (e) and (g). As will be evident, the exo-sample nucleotide may be incorporated in the amplification product during the amplification steps and as part of the primer. The two oligonucleotide primers of step (e) may or may not be identical.

In addition, the invention is directed to a process for amplifying one or more nucleic acid molecules in a sample, comprising the steps of:

(A) with a first sample containing a nucleic acid sequence or a mixture of nucleic acid molecules, wherein each nucleic acid molecule has two separate complementary strands of equal or unequal length:

(a) treating such strands with two oligonucleotide primers, for each different specific nucleic acid molecule being amplified, wherein at least one primer

comprises an exo-sample nucleotide, under conditions such that for each different nucleic acid molecule being amplified an extension product of each primer is synthesized which is complementary to each nucleic acid strand, wherein such primers are selected so as to be sufficiently complementary to different strands of each specific nucleic acid molecule to hybridize therewith such that the extension product synthesized from one primer, when it is separated from its complement, can serve as a template for synthesis of the extension product of the other primer;

(b) separating the primer extension products from the nucleic acid molecule templates on which they were synthesized to produce single-stranded molecules;

(c) treating such single-stranded molecules generated from step (b) with the primers of step (a) under conditions such that a primer extension product is synthesized using each of the single strands produced in step (b) as a template;

(d) repeating steps (a) through (c) at least once, thereby amplifing the specific nucleic acid sequence contained by such first sample; and

(B) with a second sample containing a nucleic acid molecule or a mixture of nucleic acid molecules, wherein each nucleic acid molecule has two separate complementary strands of equal or unequal length and wherein amplified nucleic acid molecule of step (A) may be present in such second sample:

(e) subjecting the nucleic acid of a second sample to treatment which renders nucleic acid containing such exo-sample nucleotide substantially unamplifiable in an amplification dependent on such specific primers and which does not substantially affect amplification of nucleic acid that does not contain the exo-sample nucleotide;

whereby any primer extension products of such first sample, amplified in step (A) and present in such second sample, are not further substantially amplified in step (B).

Most particularly, the present invention is directed to a process for amplifying one or more nucleic acid molecules in a sample, comprising the steps of:

(A) with a first sample containing a nucleic acid molecule or a mixture of nucleic acid molecules, wherein each nucleic acid molecule has two separate complementary strands of equal or unequal length:

(a) treating such strands with two oligonucleotide primers, for each different specific sequence being amplified, wherein at least one primer comprises deoxyuridine, under conditions such that for each different nucleic acid molecule being amplified an extension product of each primer is synthesized which is complementary to each nucleic acid strand, wherein such primers are selected so as to be sufficiently complementary to different strands of each specific nucleic acid molecule to hybridize therewith such that the extension product synthesized from one primer, when it is separated from its complement, can serve as a template for synthesis of the extension product of the other primer;

(b) separating the primer extension products from the templates on which they were synthesized to produce single-stranded molecules;

(c) treating such single-stranded molecules generated from step (b) with the primers of step (a) under conditions that a primer extension product is synthesized using each of the single strands produced in step (b) as a template;

(d) repeating steps (a) through (c) at least once, thereby amplifying such specific nucleic acid molecule contained by the first sample;

(B) with a second sample containing a nucleic acid sequence or a mixture of nucleic acid sequences, wherein each nucleic acid sequence has two separate complementary strands of equal or unequal length and wherein amplified nucleic acid sequences of step (A) may be present in such second sample:

(e) treating such strands with uracil DNA glycosylase;

(f) terminating the action on the strands of the uracil DNA glycosylase by heating; and

(g) repeating steps (a) through (c) at least once, thereby amplifying any of the specific nucleic acid sequence contained by second sample;

whereby any primer extension products of the first sample, amplified in steps (A) and present in such second sample, are not further substantially amplified in steps (B).

In addition, the present invention is directed to a method of detecting target nucleic acid in a sample comprising the steps of:

(a) providing nucleic acid of the sample as single-stranded nucleic acid;

(b) providing in the sample at least four DNA probes, wherein:

i) the first and second of said probes are primary probes, and the third and fourth of said probes are secondary nucleic acid probes;

ii) the first probe is a single strand capable of hybridizing to a first segment of a primary strand of the target nucleic acid;

iii) the second probe is a single strand capable of hybridizing to a second segment of said primary strand of the target nucleic acid sequence;

iv) the 5' end of the first segment of said primary strand of the target is positioned relative to the 3' end of the second segment of said primary strand of the target to enable joining of the 3' end of the first probe to the 5' end of the second probe, when said probes are hybridized to said primary strand of said target nucleic acid;

v) the third probe is capable of hybridizing to the first probe;

vi) the fourth probe is capable of hybridizing to the second probe;

vii) at least the 3' nucleotide of the first probe or the 5' nucleotide of the second probe is deoxyuridine; and

viii) at least the 3' nucleotide of the fourth probe or the 5' nucleotide of the third probe is deoxyuridine; and

(c) repeatedly performing the following cycle:

i) hybridizing said probes with nucleic acid in said sample;

ii) ligating hybridized probes to form reorganized fused probe sequences; and

iii) denaturing DNA in said sample; and

(d) detecting the reorganized fused probe sequences; whereby with successive cycles the quantity of reorganized fused primary and fused secondary probes is increased.

The aforementioned method may also comprise, according to the invention, an additional step of subjecting, before step (c), the sample to a treatment which renders amplified nucleic acid containing exo-sample nucleotide substantially unamplifiable in an amplification dependent on such specific

probes and which does not substantially affect amplification of nucleic acid that does not contain the exo-sample nucleotide.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** shows a schematic of the application of exo-sample nucleotides in ligase chain reaction.

DEFINITIONS

In the description that follows, a number of terms used in molecular biology and nucleic acid amplification technology are extensively utilized. In order to provide a clearer and consistent understanding of the specification and claims, including the scope to be given such terms, the following definitions are provided.

"Amplification", as used herein, refers to any in vitro process for increasing the number of copies of a nucleotide sequence or sequences. Nucleic acid amplification results in the incorporation of nucleotides into DNA or RNA. As used herein, one amplification reaction may consist of many rounds of DNA replication. For example, one PCR reaction may consist of 30–100 "cycles" of denaturation and replication.

"Nucleotide" as used herein, is a term of art that refers to a base-sugar-phosphate combination. Nucleotides are the monomeric units of nucleic acid polymers, i.e. of DNA and RNA. The term includes ribonucleoside triphosphates, such as rATP, rCTP, rGTP, or rUTP, and deoxyribonucleotide triphosphates, such as dATP, dCTP, dUTP, dGTP, or dTTP. A "nucleoside" is a base-sugar combination, i.e. a nucleotide lacking phosphate. It is recognized in the art that there is a certain inter-changability in usage of the terms nucleoside and nucleotide. For example, the nucleotide deoxyuridine triphosphate, DUTP, is a deoxyribonucleoside triphosphate. After incorporation into DNA, it serves as a DNA monomer, formally being deoxyuridylate, i.e. dUMP or deoxyuridine monophosphate. One may say that one incorporates dUTP into DNA even though there is no dUTP moiety in the resultant DNA. Similarly, one may say that one incorporates deoxyuridine into DNA even though that is only a part of the substrate molecule.

"Exo-sample nucleotide", as used herein, refers to a nucleotide which is generally not found in the sample or sequence to be amplified. For most DNA samples, deoxyuridine is an example of an exo-sample nucleotide. Although the triphosphate form of deoxyuridine, dUTP, is present in living organisms as a metabolic intermediate, it is rarely incorporated into DNA. When dUTP is incorporated into DNA, the resulting deoxyuridine is promptly removed in vivo by normal processes, e.g. processes involving the enzyme uracil DNA glycosylase (UDG). Thus, deoxyuridine occurs rarely or never in natural DNA. It is recognized that some organisms may naturally incorporate deoxyuridine into DNA. For nucleic acid samples of those organisms, deoxyuridine would not be considered an exo-sample nucleotide. The presence of deoxyuridine, or any other exo-sample nucleotide, may be determined readily using methods well known to the art. Other exo-sample nucleotides may be envisioned. Numerous DNA glycosylases are known to the art. Any exo-sample nucleotide which may be incorporated into DNA during an amplification and a DNA glycosylase that acts on it may be used in this invention.

"Uracil DNA glycosylase" (UDG), a term of art, refers to an enzyme which cleaves the glycosidic bond between the base uracil and the sugar deoxyribose, only when the monomeric nucleotide dUTP is incorporated into a DNA molecule, resulting in incorporation of a deoxyuridine moiety (Duncan, B. in *The Enzymes* 14:565 (1981), ed.: Boyer P). The enzyme does not act upon free dUTP, free deoxyuridine, or RNA (Duncan, supra).

"Incorporating" as used herein, means becoming part of a nucleic acid polymer.

"Terminating" as used herein, means causing a treatment to stop. The term includes means for both permanent and conditional stoppages. For example, if the treatment is enzymatic, a permanent stoppage would be beat denaturation; a conditional stoppage would be, for example, use of a temperature outside the enzyme's active range. Both types of termination are intended to fall within the scope of this term.

"Oligonucleotide" as used herein refers collectively and interchangeably to two terms of art, "oligonucleotide" and "polynucleotide". Note that although oligonucleotide and polynucleotide are distinct terms of art, there is no exact dividing line between them and they are used interchangeably herein. The term "probe" may also be used interchangeably with the terms "oligonucleotide" and "polynucleotide".

"Oligonucleotide-dependent amplification" as used herein refers to amplification using an oligonucleotide or polynucleotide or probe to amplify a nucleic acid sequence. An oligonucleotide-dependent amplification is any amplification that requires the presence of one or more oligonucleotides or polynucleotides or probes that are two or more mononucleotide subunits in length and that end up as part of the newly-formed, amplified nucleic acid molecule.

"Primer" as used herein refers to a single-stranded oligonucleotide or a single-stranded polynucleotide that is extended by covalent addition of nucleotide monomers during amplification. Nucleic acid amplification often is based on nucleic acid synthesis by a nucleic acid polymerase. Many such polymerases require the presence of a primer that can be extended to initiate such nucleic acid synthesis.

"Substantially unamplifiable" as used herein refers to the inhibition of amplification of nucleic acids in any in vitro process. When the nucleic acid is said to be "substantially unamplifiable", the majority of the nucleic acid molecules ($\geqq$50%) cannot be or are inhibited from being amplified, for example, by PCR or LCR. Although the term "substantially unamplifiable" indicates that the amplification of nucleic acid sequences is prevented or inhibited from being amplified, some amplification may still occur but such "background" amplification of a "substantially unamplifiable" nucleic acid can be distinguished from amplification of other nucleic acids which are "substantially amplifiable." Nucleic acids which are "substantially amplifiable" are, relatively speaking, amplified to a greater extent (at least two fold more) than "substantially unamplifiable" nucleic acids. One of ordinary skill in the art may easily utilize controls to distinguish between amplification of "substantially unamplifiable" nucleic acids and "substantially amplifiable" nucleic acids in accordance with the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

In the process according to the first embodiment, an amplification is procedure is performed on a first sample in which one or more of the four normal ribonucleoside triphosphates (rNTPs) or deoxyribonucleoside triphosphates (dNTPs) is replaced with an exo-sample nucleotide. Before a second amplification, any contaminating amplified product that may be remaining is subjected to a physical, chemical,

enzymatic, or biological treatment which renders nucleic acid containing the exo-sample nucleotide substantially unamplifiable. The treatment may be done as a separate step, or preferably, may be done in the presence of a second sample containing nucleic acid sequences to be amplified. The amplified nucleic acid sequences derived from the first sample which contaminate the second sample are not further substantially amplified during amplification of nucleic acid sequences of the second sample.

The deoxyribonucleoside triphosphate dUTP exemplifies an exo-sample nucleotide which may be conveniently incorporated into an enzymatic DNA amplification procedure, exemplified herein by PCR, thereby resulting in deoxyuridine-containing DNA. The DNA products of such a reaction will normally contain many uracil bases. Discrimination between natural DNA and the resultant, deoxyuridine-containing products of amplification procedures may be obtained with the enzyme uracil DNA glycosylase (UDG). Treatment of DNA containing uracil bases with uracil DNA glycosylase results in cleavage of the glycosidic bond between the deoxyribose of the DNA sugar-phosphate backbone and the uracil base. The loss of the uracil creates an apyrimidinic site in the DNA, which blocks DNA polymerase from using the DNA strand as a template for the synthesis of a complementary DNA strand (Schaaper, R., et al., *Proc. Natl. Acad. Sci. USA* 80:487 (1983)). The presence of substantial numbers of apyrimidinic sites in each DNA target molecule interferes with amplification procedures which use DNA polymerase to synthesize copies of target DNA. The DNA sugar-phosphate backbone that remains after UDG cleavage of the glycosidic bond can then be cleaved by endonuclease IV, alkaline hydrolysis, tripeptides containing aromatic residues between basic ones such as Lys-Trp-Lys and Lys-Tyr-Lys (Pierre et al., *J. Biol Chem.* 256:10217–10226 (1981)) and the like.

As exemplified herein, the basic amplification protocol is the well known PCR method. PCR was modified in three ways: (1) dUTP was substituted for dTTP; (2) UDG was added to the initial PCR reaction mixture; and (3) an initial incubation period was added to allow UDG to destroy contaminating products of prior PCR reactions. The UDG itself was either permanently inactivated by high temperature in the first PCR cycle or was not active at the high temperatures used with Taq polymerase in the currently preferred PCR protocol. This inactivation prevents UDG from destroying newly-synthesized PCR products. Nucleic acid amplification protocols that do not eliminate UDG activity usually will require an extra UDG-inactivation step.

While termination of the physical, chemical, enzymatic, or biological treatment that renders nucleic acid containing the exo-sample nucleotide resistant to the amplification process is preferred (as exemplified herein, heat inactivation of UDG), the embodiment also includes variations lacking a termination step. For example, one might use amounts of enzyme and durations of treatment high enough to eliminate expected contamination of starting materials but insufficient to keep up with the rate of amplification. In other words, a treatment might be able to destroy contaminating nucleic acid but an amplification process might still be able to produce new nucleic acid faster than the treatment could destroy the newly synthesized nucleic acid.

Variations on the herein disclosed embodiment may also be envisioned. For example, the amplification may be done without an exo-sample nucleotide, i.e., using normal nucleotides. A normal nucleotide in the amplified DNA is then converted into an exo-sample nucleotide. The converted DNA can then be removed from any samples which it later contaminates. An example would be the conversion of neighboring pyrimidine residues, especially thymidine, into pyrimidine dimers (thymidine dimers), which make DNA unsuitable as a template. Thymidine dimers can also be removed by enzymes such as exonuclease VII and recBC.

The methods of the invention allow amplification products to be distinguished from the original nucleic acid substrate in a manner which does not interfere with detection analysis of such amplification products but which does make such amplification products uniquely susceptible to certain treatments which destroy their ability to serve as templates in subsequent amplification reactions. Because of this susceptibility, pretreating samples to be amplified according to the methods of this embodiment allows the elimination of any contaminating products of previous amplifications from the sample.

In the first embodiment, an exo-sample nucleotide is incorporated into amplification products during the amplification procedure itself so as to distinguish such amplified sequences from the original template.

In the second embodiment, the exo-sample nucleotide used to distinguish such amplification products herein is provided as part of an oligomer or polymer before amplification, while that of the first embodiment is provided as a nucleoside triphosphate which becomes incorporated into polymeric nucleic acid during amplification.

The methods of the second embodiment are applicable to any in vitro procedures which utilize enzymes to amplify specific nucleic acid sequences and especially to PCR and LCR.

In a preferred variation of the second embodiment, the nucleic acid products produced by an amplification which uses primers containing exo-sample nucleotides are chemically different from starting templates which were not produced by an amplification containing exo-sample nucleotides. This chemical difference permits one to render nucleic acid products, which were produced by an amplification which used primers containing exo-sample nucleotides, incapable of further exponential amplification.

Incorporation of an exo-sample nucleotide into a primer or probe allows the DNA or RNA produced during such amplification processes to be differentiated from the original nucleic acids present in the sample prior to amplification. If desired, the amplification reaction itself may, in addition, further provide exo-sample nucleotides for incorporation into the replicating nucleic acid.

Typically, oligonucleotides are used wherein one or more of the four ribonucleotides (ATP, UTP, CTP and GTP), or deoxyribonucleotides (dATP, dTTP, dCTP and dGTP), in the oligonucleotide are replaced with one or more exo-sample nucleotides. Embodiments utilizing oligonucleotides with high proportions of exo-sample nucleotides are preferred over those with fewer exo-sample nucleotide-containing oligonucleotides. Especially, oligonucleotides with a high fraction of exo-sample nucleotides located at the 3'OH end of the oligonucleotide are preferred. In another preferred embodiment, an exo-sample nucleotide is the 3' nucleotide.

In one variation of the second embodiment, treatment of exo-sample nucleotide primers to render amplified nucleic acid unamplifiable is performed before beginning an amplification reaction.

In another variation of the second embodiment, treatment to render amplified nucleic acid un-amplifiable is performed after terminating a first amplification reaction and after removing a sample of the amplified products for further analysis, but before subjecting the original, starting nucleic

acid in the sample to a new, second amplification reaction. For example, after amplification and after removing amplified samples for analysis, amplified product that may be remaining is subjected to a treatment which renders amplified nucleic acid containing the exo-sample nucleotide substantially unamplifiable.

Such treatment to render the amplified nucleic acid containing the exo-sample nucleotide unamplifiable may be a physical, chemical, enzymatic, or biological treatment. Such treatment may be done as a separate step, or preferably, may be done in the presence of a second sample containing nucleic acid sequences to be amplified. Accordingly, amplified nucleic acid sequences produced by amplification of a first sample in the presence of exo-sample nucleotides which contaminate a second sample are not further substantially amplified during the second oligonucleotide-dependent amplification of nucleic acid sequences even if the same oligonucleotide(s) are used as primers or probes.

Discrimination between a nucleic acid which does not contain the exo-sample nucleotide deoxyuridine and a deoxyuridine-containing product of an amplification reaction may be obtained with the enzyme UDG. Treatment of DNA containing uracil bases with UDG results in cleavage of the glycosidic bond between the deoxyribose of the DNA sugar-phosphate backbone and the uracil base. The loss of the uracil creates an apyrimidinic site in the DNA, which blocks DNA polymerase from using the DNA strand as a template for the synthesis of a complementary DNA strand (Schaaper, R. et al. *Proc. Natl. Acad. Sci. USA* 80:487 (1983)). The presence of substantial numbers of apyrimidinic sites in each DNA target molecule interferes with amplification procedures which use DNA polymerase to synthesize copies of target DNA. The DNA sugar-phosphate backbone that remains after UDG cleavage of the glycosidic bond can then be cleaved by endonuclease IV, alkaline hydrolysis, tripeptides containing aromatic residues between basic ones such as Lys-Trp-Lys and Lys-Tyr-Lys (Pierre et al., *J. Biol. Chem.* 256:10217–10226 (1981) and Doetsch et al. *Mutation Research* 236:173–201 (1990)) and the like.

By providing primers containing exo-sample nucleotides, such exo-sample nucleotides are localized at the 5' ends of each strand of DNA template which is amplified. When deoxyuridine-containing primers have been used and the sample treated with UDG, substantial numbers of apyrimidinic sites in the 5'-end of each DNA target template molecule are found. Such apyrimidinic sites interfere with extension at the 3'-end of newly made strands. These 3'-end sequences are the targets to which original exo-sample-containing primers bind. Thus these primers do have target sequences with which they can bind a nucleic acid derived from molecules primed by exo-sample nucleotide-containing primers.

Oligonucleotides that contain deoxyuridine at or near their termini are not substantially susceptible to UDG cleavage of the glycosidic bond between the deoxyribose of the DNA sugar-phosphate backbone and the uracil base. Thus, particularly in LCR reactions, such oligonucleotides containing deoxyuridine at or near their termini are amplifiable even after treatment with UDG.

After amplification, preferably by LCR, the joined or ligated oligonucleotides (probes) containing deoxyuridine are substantially cleaved by UDG, thereby forming an apyrimidinic deoxyribose in a sugar-phosphate backbone of the ligated or joined oligonucleotides. The resulting sugar-phosphate backbone can be cleaved by treatment with endonuclease IV, alkaline hydrolysis, tripeptides containing aromatic residues between basic ones (Pierre et al., *J. Biol. Chem.* 256:10271–10220 (1991)) and the like. Thus, UDG treatment, and optionally the cleavage of the sugar-phosphate backbone, substantially prevents amplification of the ligation product, thereby allowing one to distinguish and eliminate such amplification products as possible contaminants from subsequent amplification reactions.

Other exo-sample nucleotides than deoxyuridine, exemplified herein, may be envisioned. Numerous DNA glycosylases are known to the art. An exo-sample nucleotide which may be chemically or enzymatically incorporated into an oligonucleotide and a DNA glycosylase that acts on it may be used in this invention. Other exo-sample nucleotides will be evident to those in the art. For example, RNA primers used for DNA amplifications can be readily destroyed by base or an appropriate ribonuclease (RNase). RNase H degrades RNA of RNA:DNA hybrids and numerous single-stranded RNases are known which are useful to digest single-stranded RNA after a denaturation step. (Note that unless the DNA strand opposite the RNA primer is removed by a single-stranded deoxyribonuclease (DNase) after RNase H digestion of the contaminating DNA but before denaturation thereof, the contaminating DNA will be amplified in a linear fashion. However, since most oligonucleotide-dependent DNA amplifications schemes amplify DNA exponentially, this will not be a problem unless the contamination levels are extraordinarily high. Those of ordinary skill in the art who are developing RNA-primed variations of the invention will be able to determine, without undue experimentation, the lowest acceptable level of contamination in their assays.)

In another embodiment deoxyoligonucleotides containing bromodeoxyuridine (BdUR) is used as the exo-sample nucleotide. DNA containing BdUR may be degraded on exposure to light under appropriate conditions.

As exemplified herein, in a preferred variation of the second embodiment, the basic amplification protocol PCR is modified in three ways: (1) amplification is first performed using oligonucleotide primers containing deoxyuracil substituted for deoxythymidine and samples of the amplified products are removed for analysis; (2) UDG is then added to subsequent PCR reaction mixtures; and (3) an incubation period is added to allow UDG to act on uracil-containing sequences in contaminating products from prior PCR reactions.

As exemplified herein, in another preferred variation of the second embodiment, the basic amplification protocol LCR is modified in three ways: (1) amplification is first performed using oligonucleotide probes containing deoxyuracil substituted for deoxythymidine at or near at least one terminus of the oligonucleotides and samples of the amplified products are removed for analysis; (2) UDG is then added to subsequent LCR reaction mixtures; and (3) an incubation period is added to allow UDG to act on uracil-containing sequences in contaminating products from prior LCR reactions. The UDG treatment alone has been found to substantially prevent amplification of the uracil-containing ligation product without substantially affecting amplification of the oligonucleotide probes containing deoxyuracil near the end of the probes. As an alternative to UDG treatment alone, UDG and a treatment that cleaves the sugar-phosphate backbone (remaining after UDG treatment) may be used. Such treatments that cleave the sugar-phosphate backbone may include, but are not limited to, endonuclease IV, alkaline hydrolysis, tripeptides such as Lys-Trp-Lys and Lys-Tyr-Lys (Pierre et al., *J. Biol. Chem.* 256:10217–10220 (1981)), AP endonucleases such as endonuclease V, endo-

nuclease III, endonuclease VI, endonuclease VII, human endonuclease II, and the like (Linn, S. *Nucleases Involved in DNA Repair* in *Nucleases*, Cold Spring Harbor Laboratory, ed. Linn, S. and Roberts, R. (1985)).

Before continuing with the amplification procedure, the UDG and the enzyme treatment to cleave the sugar-phosphate backbone (if present) should be inactivated. UDG and the enzyme treatment to cleave the sugar-phosphate backbone may be either permanently inactivated by high temperature in the first PCR or LCR cycle, or by the high temperatures used with Taq DNA polymerase or Taq DNA ligase in the currently preferred PCR or LCR protocols, respectively. This inactivation prevents UDG from destroying newly-synthesized PCR or LCR products. Nucleic acid amplification protocols that do not inherently eliminate UDG activity usually will require an extra UDG-inactivation step.

Embodiments having a greater proportion of oligonucleotides containing exo-sample nucleotides are preferred. However, even in standard PCR embodiments that depend on two oligonucleotide primers, the present invention is capable of rendering PCR contaminants unamplifiable as long as at least one primer contains an exo-sample nucleotide.

Not all single exo-sample nucleotide-containing oligonucleotides may reduce amplification of PCR or LCR products to sufficiently low levels for a given situation; those of ordinary skill in the art can empirically find which primers or probes are acceptable during routine optimization and testing of a PCR or LCR assay, without undue experimentation.

Routine assay optimization, aimed at testing oligonucleotide suitability, can be done by (1) making an exo-sample nucleotide-containing oligonucleotide, (2) performing a first nucleic acid amplification of a target sequence using that oligonucleotide, (3) seeding various amounts of the resulting first product in a new, second amplification that does not contain target sequences, (4) treating the second amplification to render exo-sample nucleotide-containing nucleic acid unamplifiable, (5) performing the second amplification, and (6) assaying the resultant second product for presence of contaminating sequence from the first hybridization. All of these steps are routinely done as experimental controls and as part of the normal validation of a method. The only additional work required of one practicing this embodiment involves routine synthesis of any additional oligonucleotides that may be required. Generally, if a oligonucleotide is observed to be not suitable, a substitute can be easily found and similarly tested.

As will be understood by those of ordinary skill in the art, the methods of the invention prevent exponential amplification but may not prevent linear amplification. Linear amplification does not usually represent a substantial problem in amplification procedures. For example, a single amplification product molecule which contaminates a sample run through 20 cycles of PCR, amplifying by a factor of two on each cycle, will result in only about 20 molecules if amplified linearly but could result in up to about a million molecules if amplified exponentially at maximal theoretical efficiency. Thus, linear amplification is generally inconsequential.

The second embodiment can be adapted to be used with amplifications that do not use a nucleic acid polymerase. For instance, as exemplified herein, contamination of samples with previously made products can be controlled by amplification schemes that polymerize oligonucleotides with ligase (e.g. Wu and Wallace, supra, and Barany, *Proc. Natl. Acad. Sci. USA* 88:189–193 (1991)).

The present invention is also ideally suited for preparation of a kit. Such a kit may comprise a carrier means being compartmentalized to receive in close confinement one or more container means such as vials, test tubes, bottles, and the like. Each of said container means may comprise an oligonucleotide containing exo-sample nucleotide such as an oligonucleotide containing deoxyuridine, an exo-sample nucleotide such as deoxyuridine, a nucleic acid ligase, a nucleic acid polymerase, a uracil DNA glycosylase and the like. As will be evident, the components included in a kit such as a nucleic acid ligase, a uracil DNA glycosylase etc. may be mixed in various combinations and such combinations may be included in a single or multiple container means.

Having now generally described this invention, the same will be better understood by reference to certain examples which are included herein for purposes of illustration only and is not intended to be limiting unless otherwise specified.

EXAMPLES

Example 1

A polymerase chain reaction (PCR) was performed to amplify a region of the human papilloma virus type 16 (HPV 16) DNA (Durst, M. et al., *Proc. Natl. Acad. Sci. USA* 80:3812 (1983)). The sequences of the primers used were 5'GGTCGATGTATGTCTTGTTG3' and 5'GTCTACGTGTGTGCTTTGTAC3'.

HPV 16 DNA was excised from a full length plasmid clone, pT7HPV16 (for the purposes of this invention, equivalent to the pUC8 plasmids described by Seedoff, K., et al., *Virol.* 145:181 (1985)) with the restriction enzyme BamH I. The linear DNA (10 picograms) was added to PCR reactions containing 50 microliters of 25 mM Tris HCl pH 8.3, 5 mM $MgCl_2$, 50 mM NaCl, 0.01% gelatin, 0.05% W1 polyoxyethylene ether detergent (Sigma), 0.2 mM each dATP, dGTP, dCTP, 0.2 mM either dUTP or dTTP, 1 micromolar of each primer, and 2.5 units of thermostable DNA polymerase from *Thermus aquaticus* (Cetus/Perkin-Elmer). The reactions were amplified in a thermal cylcer (Cetus/Perkin-Elmer) using the following temperature profile: 5 minutes at 94° C., then 30 cycles of 1 minute at 94° C. (denaturation), two minutes at 55° C. (annealing), and 3 minutes at 72° C. (primer extension). After completion of the temperature cycles, a final extension of 10 minutes at 72° C. was done. Amplification of the 284 base pair HPV 16 DNA fragment was confirmed by agarose/ethidium bromide gel electrophoresis (Maniatis, T., et al., *Molecular Cloning*, Cold Spring Harbor Laboratory (1982)) of the PCR reaction products (5 microliters of each reaction per lane). All reactions showed substantial amplification. Negative control reactions to which no HPV 16 DNA was added did not produce any visible DNA products.

The concentration of the PCR amplification products was estimated from the agarose gel. New PCR reactions were contaminated with ten femtogram quantities of the amplification products that contained either deoxythymidine, resulting from incorporation of dTTP, or deoxyuridine, from dUTP-containing reactions. Positive control reactions contained 10 picograms of linear HPV 16 DNA. Negative control reactions did not receive any target DNA. The new PCR reactions contained dUTP instead of dTTP, and either 5 nanograms of UDG (Van de Sande, H., University of Calgary; also available from Life Technologies Inc., P.O.

Box 9418, Gaithersburg, Md., 20898) or no UDG. All reactions were incubated for 15 minutes at 37° C. to allow the UDG to act on deoxyuridine-containing DNA, and then were taken through the same thermal cycling protocol as above. Aliquots of each reaction were analyzed by agarose/ethidium bromide gel electrophoresis.

The agarose gel analysis showed that without UDG treatment the deoxyuridine-containing PCR products could be re-amplified to give a DNA product virtually indistinguishable in size, as evidenced by gel electrophoresis, from the products obtained by amplifying the normal HPV 16 DNA. Reactions in which the deoxyuridine-containing DNA was incubated with UDG prior to PCR did not give any visible products on the agarose gel. PCR amplification products that contained deoxythymidine were successfully amplified whether or not they had been incubated with UDG. This experiment showed that UDG substantially abolished amplification of PCR products containing deoxyuridine, but had no substantial effect on the amplification of DNA containing deoxythymidine.

Although the foregoing refers to particular preferred variations of the first embodiment, it will be understood that this embodiment is not so limited. It will occur to those of ordinary skill in the art that various modifications may be made to the disclosed embodiments and that such modifications are intended to be within the scope of this embodiment.

Example 2

General Experimental Conditions of the Second Embodiment

Polymerase chain reactions (PCRs) were performed to amplify a region of the human papilloma virus type 16 (HPV 16) DNA (Durst, M. et al. *Proc. Natl. Acad. Sci. USA* 80:3812 (1983)). The sequences of the oligonucleotide primers used were either 5'GGUCGAUGUAUGUCUUGUUG3' and 5'GUCUACGUGUGUGCUUUGUAC3' (dU primers dU_1 and dU_2, respectively) or 5'GGTCGATGTATGTCTTGTTG3' and 5'GTCTACGTGTGTGCTTTGTAC3' (control dT primers dT_1 and dT_2, respectively). Note that the sequences of dT_1 and dU_1 were identical except for the replacement of U for T; dT_2 and dU_2 were similarly identical.

HPV 16 DNA was excised from a full length plasmid clone, pT7HPV16 (for the purposes of this invention, equivalent to the pUC8-based plasmids described by Seedoff, K. et al. *Virol.* 145:181 (1985)) with the restriction enzyme BamH I. The linear DNA (10 picograms) was added to PCR reactions containing 50 microliters of 25 mM Tris HCl pH 8.3, 5 mM $MgCl_2$, 50 mM NaCl, 0.01% gelatin, 0.2 mM each dATP, dGTP, dCTP, 0.2 mM either dUTP or dTTP, 1 micromolar of each primer, and 2.5 units of thermostable DNA polymerase from *Thermus aquaticus* (Cetus/Perkin-Elmer). Primers and triphosphates were paired as follows: dU primers with dTTP (pUxT reactions ("primed with U, extended with T")) and control dT primers with dUTP (pTxU reactions). The reactions were amplified in a thermal cylcer (Cetus/Perkin-Elmer) using the following temperature profile: 5 minutes at 94° C., then 30 cycles of 1 minute at 94° C. (denaturation), two minutes at 55° C. (annealing), and 3 minutes at 72° C. (primer extension). After completion of the temperature cycles, a final extension of 10 minutes at 72° C. was done. Amplification of the 284 base pair HPV 16 DNA fragment was confirmed by agarose/ethidium bromide gel electrophoresis (Maniatis, T., et al., *Molecular Cloning*, Cold Spring Harbor Laboratory (1982)) of the PCR reaction products (5 microliters of each reaction per lane). All reactions showed substantial amplification. Negative control reactions to which no HPV 16 DNA was added did not produce any DNA products. The concentration of the PCR amplification products was estimated from the agarose gel.

Example 3

Amplification Using Two Deoxyuridine-Containing Primers

New pUxT PCR reactions were prepared that contained dTTP but lacked dU primers. These were contaminated with ten femtograms of pUxT amplification products from the previous reactions, and incubated for 15 minutes at 37° C. in the presence or absence of UDG (H. Van de Sande, University of Calgary; also available from Life Technologies Inc., P.O. Box 9418, Gaithersburg, Md. 20898). These reactions were then incubated at 94° C. to inactivate the UDG, cooled to 15° C., and dU primers were added.

New pTxU PCR reactions were prepared with dT primers and dUTP and were contaminated with ten femtograms of pTxU amplification products. These reactions were incubated at 37° C. for 15 minutes in the presence or absence of UDG to act as controls for UDG activity and PCR amplification, respectively.

The new PCR reactions were then subjected to the same PCR temperature cycling described above. Agarose/ethidium bromide gel electrophoresis showed that in the absence of UDG treatment both pUxT and pTxU reaction products could be reamplified by PCR to give results identical in appearance to those obtained upon amplification of HPV 16 DNA. In contrast, UDG treatment abolished reamplification of both pTxU reaction products (the positive control) and pUxT reaction products, which contained dU primers. In another experiment it was shown that, in the absence of deoxyuridine in either the primers or extension products, UDG did not affect the PCR amplification of natural HPV 16 DNA.

Without UDG treatment, the deoxyuridine-containing PCR products could be re-amplified to give a DNA product virtually indistinguishable in size, as evidenced by agarose gel electrophoresis, from the products which did not contain deoxyuridine and which were obtained by amplifying the normal HPV 16 DNA. The reaction in which the DNA was made with deoxyuridine-containing primers and which was incubated with UDG prior to PCR did not substantially give any visible products on the agarose gel. In another experiment it was shown that UDG did not affect the amplification of natural DNA.

Example 4

Amplification Using a Single Deoxyuridine-Containing Primer

A further experiment demonstrated that only one of two PCR primers need contain an exo-sample nucleotide. PCR amplifications were done essentially as described above using the following pairs of primers: (1) dT_1 with dT_2, (2) dT_1 with dU_2, (3) dU_1 with dT_2, and (4) dU_1 with dU_2. All combinations of PCR products were treated both with and without UDG before the second round of PCR. Amplification (1) tested for success of PCR amplification while amplification (4) tested for activity of UDG and the ability to eliminate contamination. Amplification (3) was observed to not be substantially reamplified after UDG treatment in a

second round of PCR, thereby demonstrating that only one of a pair of PCR primers need contain an exo-sample nucleotide in order to reduce or eliminate the effects of sample contamination. Amplification (2) was observed to be reamplified, though at a substantially lower level than amplification (1).

Summary

In summary, PCR amplification products that contained deoxyuridine in a primer or primers were successfully amplified if no UDG was present; amplification could be prevented by prior incubation with UDG. In other words, UDG could substantially abolish amplification of PCR products made with deoxyuridine-containing primers but had no substantial effect on the amplification of DNA made with deoxythymidine-containing primers.

Example 5

A modification of the Ligase Chain Reaction (LCR) described by Barany (F. Barany, *Proc. Natl. Acad. Sci.* 88, 189–193, 1991) was performed to amplify a region of the human beta globin gene. A schematic of this experiment demonstrating the application of exo-sample nucleotides in LCR is shown in FIG. **1**.

The sequences of the probe oligonucleotides were:

1088 (probe 4): 5' acc atg gtg cac ctg act ccu 3'
1089 (probe 3): 5' gag gag aag tct gcc gtt act 3'
1090 (probe 2): 5' cag gag tca ggt gca cca tg 3'
1091 (probe 1): 5' cag taa cgg cag act tct ccu 3'

The sequence of the target oligonucleotide was:

1092 (target): 5' gac acc atg gtg cac ctg act ccu gag gag aag tct gcc gtt act gcc ctg 3'

Oligonucleotides were chemically synthesized by standard techniques. They were trace labeled with ^{32}P (approximately 30 dpm/pmol each oligonucleotide) and gel purified, and their quantities were estimated by hybridization and electrophoresis on 5% agarose/ethidium bromide gels. Thermostable DNA ligase from *Thermus aquaticus* was purchased from Molecular Biology Resources (MBR), Milwaukee, Wis. UDG was from Life Technologies, Inc., Gaithersburg, Md. The buffer used was that suggested by MBR, and contained 20 mM Tris HCl pH 7.5, 25 mM KCl, 10 mM $MgCl_2$, 10 mM dithiothreitol, 0.6 mM NAD, 0.1% Triton X100. Probe oligonucleotide #1089 was labeled at its 5' terminus with ^{32}P using T4 polynucleotide kinase to a specific activity of about 250,000 dpm per picomole, then chased with unlabelled rATP and the kinase was heat inactivated. The final concentration of the labelled #1089 oligonucleotide was about 1 picomole per microliter. Oligonucleotide #1090 was phosphorylated at its 5' end with rATP and T4 polynucleotide kinase.

Reactions for UDG treatment contained 20 pmol each of all four trace labeled probe oligos with or without 2 pmol target. These samples were incubated in 10 ul 50 mM Tris HCl pH 8.0, 1 mM EDTA under oil for 15 min at 37° in the presence or absence of 0.5 ul=0.5 units UDG, as follows:

A: Probes only, no UDG
B: Probes only, +UDG
C: Probes+target, no UDG (should support LCR)
D: Probes+target+UDG (target degraded, no LCR)
E: Probes only, +UDG

Any abasic sites (apyrimidinic sites) produced by UDG were broken (i.e., cleavage of the sugar phosphate backbone) by alkaline hydrolysis as follows: to all five tubes was added 5 ul 0.2 M NaOH giving a pH of about 13. The tubes were incubated at 95° for 10 minutes, cooled, and 5 ul of 2.5% acetic acid were added to bring the pH back to about 5. Any target (oligo # 1092) present in the reactions was cleaved by the combination of base removal by UDG and breakage of the sugar-phosphate backbone by elevated temperature and pH.

At this time (after UDG inactivation by alkaline pH and heat) 1 ul (2 pmol) target (#1092) was added to reaction E. Thus the target in reaction "E" was never exposed to active UDG, but only to inactivated UDG.

The concentrations of the salt produced by the alkaline hydrolysis reactions were reduced to avoid interference with subsequent ligase reactions and gel electrophoresis. Five ul of each reaction were diluted 1:40 in 1× MBR buffer (20 mM Tris HCl pH 7.5, 25 mM KCl, 10 mM $MgCl_2$, 10 mM dithiothreitol, 0.6 mM NAD, 0.1% Triton X100). To 20 ul each diluted oligo mixture was added 0.5 ul=500 fmol ^{32}P-labelled #1089 to follow the progress of the LCR reactions. Four ul of these mixtures containing 100 fmol probes 5'OH-1088, 5'P-1090, and 5'-OH-1091, 200 fmol 5'-^{32}P 1089, with or without 10 fmol 5'OH-target #1092, were aliquotted to pairs of tubes and 16 ul of 1× MBR buffer with ("+") or without ("−") 10 units of MBR Taq ligase were added. A drop of oil was added to the 20 ul reactions and they were cycled 85° 30s, 50° 2 min, 20 times in a Cetus/Perkin Elmer Thermal Cycler. 2.5 ul each reaction were applied to a 12% denaturing sequencing gel using standard techniques (Maniatis, supra).

Following electrophoresis and exposure of the gel to X-ray film, the following results in the presence of ligase were observed:

The "A" reaction (probes only, no UDG treatment) showed a minimal amount (about 5% of the starting probes) of target-independent amplification (i.e., ligation product of about 40 nucleotides was produced).

The "B" reaction (probes treated with UDG) showed minimal target-independent amplification indistinguishable from reaction "A". This indicated that UDG had no effect on the integrity of the probes.

The "C" reaction (containing probes and target without UDG treatment) showed the majority (about 70%) of the probes were converted to product by target-dependent ligation.

Reaction "D" (containing probes and target treated with UDG) showed substantial decrease in the amount of ligation product present. This was the expected result if the uracil in the target had been removed by UDG and the sugar-phosphate backbone had been broken by elevated pH and temperature treatment (FIG. **1**).

Reaction "E" (probes treated with UDG, and target added after UDG inactivation by pH and heat) showed the majority of the probes (about 70%) had been ligated to form product. This result demonstrated that the probes had not been affected by the UDG treatment, allowing them to participate in the ligation reactions.

The amounts of product produced in reactions C, D, and E show that UDG is active on a uracil positioned in the sequence of target oligonucleotide (i.e. not at the terminus), but is inactive on uracils at the 3' termini of oligonucleotides. It has also been demonstrated that UDG had relatively little activity on uracils located at the 5' terminus of an oligonucleotide.

To confirm the presence of uracil at the 3' terminus of probe # 1088, products of reactions C and E were pooled, precipitated with ethanol, and dissolved in 20 ul of water. One ul of 1M Tris HCl pH 8.0 and one ul of 20 mM EDTA were added to the pooled oligonucleotides and two 5 ul aliquots were placed in tubes. To one tube, 0.5 ul (0.5 units)

BRL UDG were added. To the other was added 0.5 ul water. Both tubes were incubated at 37° for 15 minutes, and to each was added 1 ul 1 N NaOH. The tubes were incubated at 95° for 15 minutes and 1 ul 12.5% acetic acid was added to neutralize the mixtures. About one third of each reaction was applied to a 12% sequencing gel. The presence of the 3' dU of oligonucleotide #1088 was confirmed by the UDG-specific cleavage of the LCR products. The size of the cleavage product was precisely that expected from the structures of the participating oligonucleotides. In other words, the ligation product of oligonucleotides 1088 and 1089 was cleaved by the combination of UDG and elevated pH and temperature, indicating that a uracil was present at the ligation junction of the participating oligonucleotides.

Summary

Uracil bases located at certain positions within oligodeoxynucleotides, especially at or near the 5' and 3' ends, were not removed by the enzyme uracil DNA glycosylase. This can be the basis of a method to control carryover contamination of Ligation Chain Reaction amplifications, if the uracil bases at these certain locations become sensitive to UDG when the oligodeoxynucleotide(s) become incorporated into LCR products.

In the example described here, deoxyuracil bases at the 3' ends of oligodeoxynucleotides are not recognized by UDG. Once the oligonucleotides containing the 3' deoxyuracil was joined or ligated by LCR, the ligation products became sensitive to cleavage with UDG and alkaline hydrolysis. Thus, this treatment rendered the amplification products unamplifiable in subsequent LCR reactions.

Example 6

A modification of LCR as described in Example 5 was performed to determine the ability of UDG alone (without alkaline hydrolysis of the sugar-phosphate backbone) to prevent amplification of uracil-containing amplification products.

20 ul reactions contained as indicated 100 fmol probe oligos (1088 and 1091 with 5' OH and 3' dU; 1089 and 1090 with 5' phosphate; 1089 labelled with ^{32}P); 5 fmol target #1092 (internal dU) or #1058 (no dU); 5 units Taq ligase; and 0.25 ul BRL UDG.

Reactions were assembled and incubated under mineral oil in a Cetus/Perkin-Elmer Thermal Cycler: 37° C. 90 min, 85° C. 5 min, then 20 cycles of 85° C. 30 sec, 55° C. 2 min, followed by incubation at 0° C.

1.25 ul aliquots of each reaction were applied to a 12.5% polyacrylamide/8M urea gel and electrophoresed at 70 watts power for about 90 minutes. An autoradiograph was made of the wet gel at −70° C. using an intensifying screen.

In the absence of target, about 5% of the probe oligo was ligated to larger product. This represents target-independent background. In the presence of 5 fmol of target, about 70 fmol of probe oligo (70% of the total present in the reaction) ligated to form product. When 0.25 ul of UDG were present in the reaction containing dT target, no effect was observed on the amount of ligation product obtained (about 70 fmol). In contrast, only about 10 fmol of ligation product was obtained when UDG was incubated with dU target (mock contaminant). These observations show that the 3' dU of oligos 1088 and 1091 were not affected by the UDG, and that greater than 90% of the dU mock contaminant was converted by UDG incubation to a form that could not be a substrate for ligation of probe oligonucleotides.

Summary

UDG treatment of uracil-containing "ligation products" (#1092) were substantially prevented from being amplified during Ligase Chain Reaction. However, uracil-containing "ligation products" not treated with UDG were amplified. Thus, UDG incubation alone, without further treatment, is sufficient to substantially prevent amplification of uracil-containing ligation contaminants which may be present in subsequent LCR reaction samples.

All references cited herein are incorporated by reference. Although the foregoing refers to particular preferred embodiments, it will be understood that the present invention is not so limited. It will occur to those of ordinary skill in the art that various modifications may be made to the disclosed embodiments and that such modifications are intended to be within the scope of the present invention.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(iii) NUMBER OF SEQUENCES: 9

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGTCGATGTA TGTCTTGTTG 20

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GTCTACGTGT GTGCTTTGTA C 21

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "RNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGUCGAUGUA UGUCUUGUUG 20

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "RNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GUCUACGUGU GUGCUUUGUA C 21

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA/RNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ACCATGGTGC ACCTGACTCC U 21

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GAGGAGAAGT CTGCCGTTAC T 21

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid

-continued

(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CAGGAGTCAG GTGCACCATG 20

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA/RNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CAGTAACGGC AGACTTCTCC U 21

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "DNA/RNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GACACCATGG TGCACCTGAC TCCUGAGGAG AAGTCTGCCG TTACTGCCCT G 51

What is claimed is:

1. A method for controlling contamination in sequential nucleic acid amplification processes comprising first and second nucleic acid amplification processes to amplify nucleic acid sequences in a first and second sample, respectively, which comprises:

(a) carrying out said first nucleic acid amplification process on said nucleic acid sequence in said first sample in the presence of an exo-sample nucleotide, to thereby generate an exo-sample nucleotide-containing product of amplification; and

(b) subjecting said exo-sample nucleotide-containing product to a treatment which renders said exo-sample nucleotide-containing product substantially unamplifiable and

wherein said treatment does not substantially prevent oligonucleotide-dependent amplification of a nucleic acid that does not contain said exo-sample nucleotide prior to carrying out said second amplification process on said nucleic acid in said second sample.

2. The method of claim **1**, wherein said exo-sample nucleotide is a substrate for a DNA glycosylase.

3. The method of claim **1**, wherein the treatment prior to the second amplification process is terminated by heating.

4. The method of claim **1**, wherein said exo-sample nucleotide is bromodeoxyuridine.

5. The method of claim **4**, wherein the treatment prior to the second amplification process is terminated by exposing the sample to light.

6. A method for oligonucleotide-dependent amplification of one or more nucleic acid molecules comprising:

a) amplifying nucleic acid of a first sample, wherein said amplifying is dependent on one or more exo-sample nucleotides to thereby produce an amplified nucleic acid comprising exo-sample nucleotides; and

b) subjecting a second sample to a treatment which renders said amplified nucleic acid comprising exo-sample nucleotides substantially unamplifiable and wherein said treatment does not substantially prevent oligonucleotide-dependent amplification of a nucleic acid that does not contain said exo-sample nucleotides.

7. The method of claim **6**, wherein said treatment is selected from the group consisting of a physical, a chemical, and an enzymatic treatment.

8. The method of claim **7**, wherein said treatment is an enzymatic treatment.

9. The method of claim **6**, wherein said exo-sample nucleotide is deoxyuridine.

10. The method of claim **9**, wherein said treatment is a treatment with uracil DNA glycosylase.

11. The method of claim **6**, further comprising amplifying the nucleic acid of said second sample which does not contain said exo-sample nucleotides.

12. The method of claim **6**, further comprising terminating said treatment.

13. The method of claim **12**, wherein said termination is accomplished by heating.

14. The method of claim **6**, wherein said exo-sample nucleotide is a substrate for a DNA glycosylase.

15. The method of claim **6**, wherein said exo-sample nucleotide is bromodeoxyuridine.

16. The method of claim **12**, wherein said termination is accomplished by exposure to light.

17. A method for amplification of nucleic acid in a sample containing nucleic acids comprising exo-sample nucleotides comprising:

a) subjecting the sample to a treatment which renders nucleic acid comprising exo-sample nucleotides substantially unamplifiable; and

b) amplifying nucleic acid of said sample which does not contain said exo-sample nucleotides.

18. The method of claim **17**, wherein said amplification is dependent on one or more exo-sample nucleotides.

19. The method of claim **17**, wherein said treatment is selected from the group consisting of a physical, a chemical, and an enzymatic treatment.

20. The method of claim **19**, wherein said treatment is an enzymatic treatment.

21. The method of claim **17**, wherein said exo-sample nucleotide is deoxyuridine.

22. The method of claim **21**, wherein said treatment is a treatment with a uracil DNA glycosylase.

23. The method of claim **17**, wherein said exo-sample nucleotide is a substrate for a DNA glycosylase.

24. The method of claim **17**, wherein said exo-sample nucleotide is bromodeoxyuridine.

25. A method of amplification and treatment of nucleic acid in a sample comprising:

a) amplifying nucleic acid in a sample in the presence of exo-sample nucleotides to thereby produce nucleic acid comprising exo-sample nucleotides, wherein said exo-sample nucleotides are substrates for DNA glycosylase; and

b) treating said nucleic acid comprising said exo-sample nucleotides with a DNA glycosylase.

26. The method of claim **25**, wherein said exo-sample nucleotides are deoxyuridine.

27. The method of claim **25**, wherein said exo-sample nucleotides are bromodeoxyuridine.

28. The method of claim **25**, wherein said DNA glycosylase is a uracil DNA glycosylase.

29. The method of claim **25**, wherein said amplification is accomplished by polymerase chain reaction (PCR).

30. The method of claim **25**, wherein said amplification is accomplished in the presence of a DNA polymerase.

31. The method of claim **30**, wherein said DNA polymerase is a Taq DNA polymerase.

32. The method of claim **25**, wherein said amplification is accomplished in the presence of one or more nucleotides in addition to the exo-sample nucleotides.

33. A method of making a treated nucleic acid molecule comprising:

a) mixing a nucleic acid template with exo-sample nucleotides;

b) incubating said mixture under conditions sufficient to synthesize a nucleic acid molecule complementary to said template, wherein said synthesized nucleic acid molecule comprises said exo-sample nucleotides; and

c) treating said synthesized nucleic acid with a DNA glycosylase.

34. The method of claim **33**, wherein said mixture further comprises a DNA polymerase.

35. The method of claim **34**, wherein said DNA polymerase is a Taq DNA polymerase.

36. The method of claim **33**, wherein said exo-sample nucleotide is deoxyuridine.

37. The method of claim **36**, wherein said DNA glycosylase is a uracil DNA glycosylase.

38. The method of claim **33**, wherein said exo-sample nucleotide is bromodeoxyuridine.

39. The method of claim **33**, wherein said mixture further comprises one or more nucleotides in addition to the exo-sample nucleotides.

40. The method of claim **33**, wherein said mixture further comprises a primer.

41. A composition comprising a nucleic acid molecule comprising one or more exo-sample nucleotides and a DNA glycosylase.

42. The composition of claim **41**, wherein said DNA glycosylase is a uracil DNA glycosylase.

43. The composition of claim **41**, wherein said exo-sample nucleotide is deoxyuridine.

44. The composition of claim **41**, wherein said exo-sample nucleotide is bromodeoxyuridine.

45. The composition of claim **41**, further comprising a DNA polymerase.

46. The composition of claim **45**, wherein said DNA polymerase is a Taq DNA polymerase.

47. The composition of claim **41**, further comprising one or more nucleotides.

48. The composition of claim **41**, further comprising a primer.

49. A composition comprising a DNA polymerase and a DNA glycosylase.

50. The composition of claim **49**, wherein said DNA glycosylase is a uracil DNA glycosylase.

51. The composition of claim **49**, wherein said DNA polymerase is a Taq DNA polymerase.

52. A kit comprising a DNA polymerase and one or more exo-sample nucleotides.

53. The kit of claim **52**, further comprising a DNA glycosylase.

54. The kit of claim **52**, further comprising one or more nucleotides.

55. The kit of claim **52**, wherein said exo-sample nucleotide is deoxyuridine.

56. The kit of claim **52**, wherein said exo-sample nucleotide is bromodeoxyuridine.

57. The kit of claim **53**, wherein said DNA glycosylase is a uracil DNA glycosylase.

58. A kit comprising a DNA polymerase and a DNA glycosylase.

59. The kit of claim **58**, further comprising one or more nucleotides.

60. The kit of claim **58**, wherein said DNA polymerase is a Taq DNA polymerase.

61. The kit of claim **58**, wherein DNA glycosylase is a uracil DNA glycosylase.

* * * * *

United States Patent [19]

Mullis

[11] **Patent Number: 4,683,202**

[45] **Date of Patent: * Jul. 28, 1987**

[54] **PROCESS FOR AMPLIFYING NUCLEIC ACID SEQUENCES**

[75] Inventor: **Kary B. Mullis,** Kensington, Calif.

[73] Assignee: **Cetus Corporation,** Emeryville, Calif.

[*] Notice: The portion of the term of this patent subsequent to Jul. 28, 2004 has been disclaimed.

[21] Appl. No.: **791,308**

[22] Filed: **Oct. 25, 1985**

Related U.S. Application Data

[63] Continuation-in-part of Ser. No. 716,975, Mar. 28, 1985, abandoned.

[51] **Int. Cl.**[4] **C12P 19/34;** C12N 15/00; C12N 1/00; C07H 21/04; C07H 21/02

[52] **U.S. Cl.** **435/91;** 435/177.3; 435/317; 536/27; 536/28; 536/29; 935/17; 935/18; 935/16

[58] **Field of Search** 435/91, 172.3, 317; 536/27, 28, 29; 935/17, 18

[56] **References Cited**

PUBLICATIONS

Gaubatz et al, "Strategies for Constructing Complementary DNA for Cloning", J. Theor. Biol. 95: 679 (1982).

Caton and Robertson, *Nucleic Acids Research,* vol. 7, pp. 1445–1456 (1979).

Rossi et al., *J. Biol. Chem.,* 257, 9226–9229 (1982).

Primary Examiner—James Martinell

Attorney, Agent, or Firm—Janet E. Hasak; Albert P. Halluin

[57] **ABSTRACT**

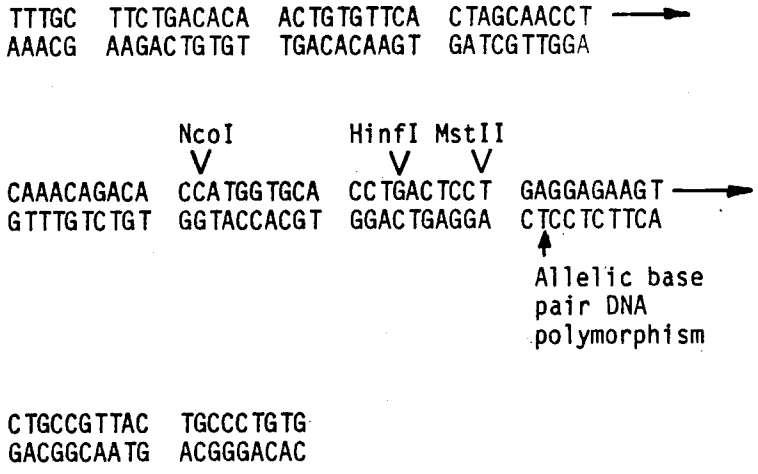
The present invention is directed to a process for amplifying any desired specific nucleic acid sequence contained in a nucleic acid or mixture thereof. The process comprises treating separate complementary strands of the nucleic acid with a molar excess of two oligonucleotide primers, and extending the primers to form complementary primer extension products which act as templates for synthesizing the desired nucleic acid sequence. The steps of the reaction may be carried out stepwise or simultaneously and can be repeated as often as desired.

21 Claims, 12 Drawing Figures

EXHIBIT F

FIG.1

Double-Stranded 94-bp Sequence



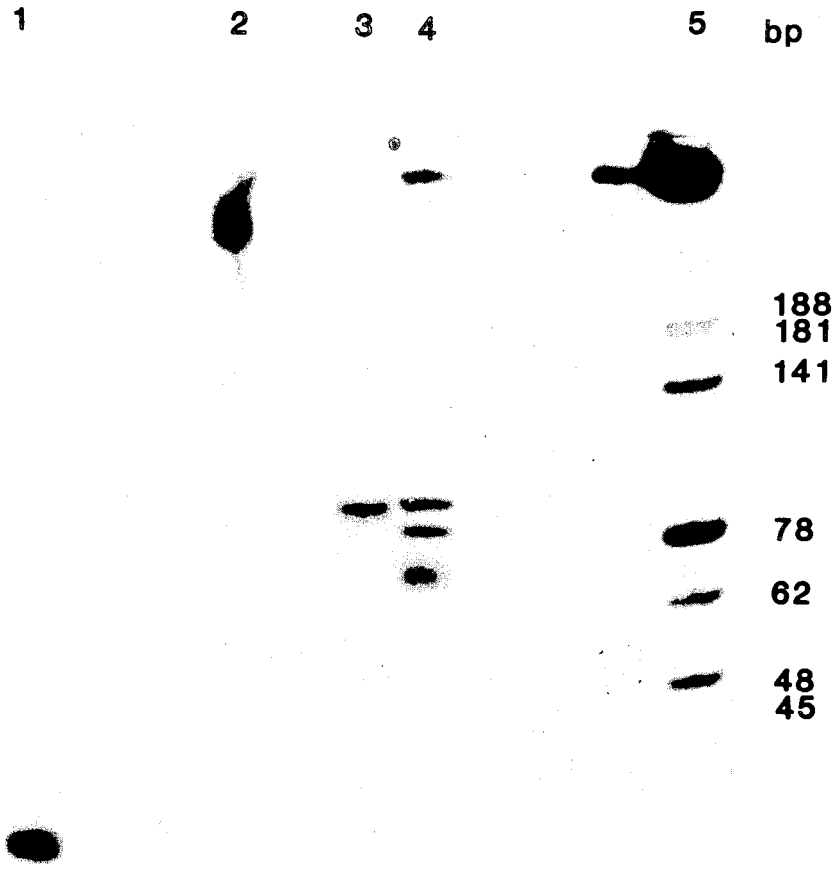


FIG.2

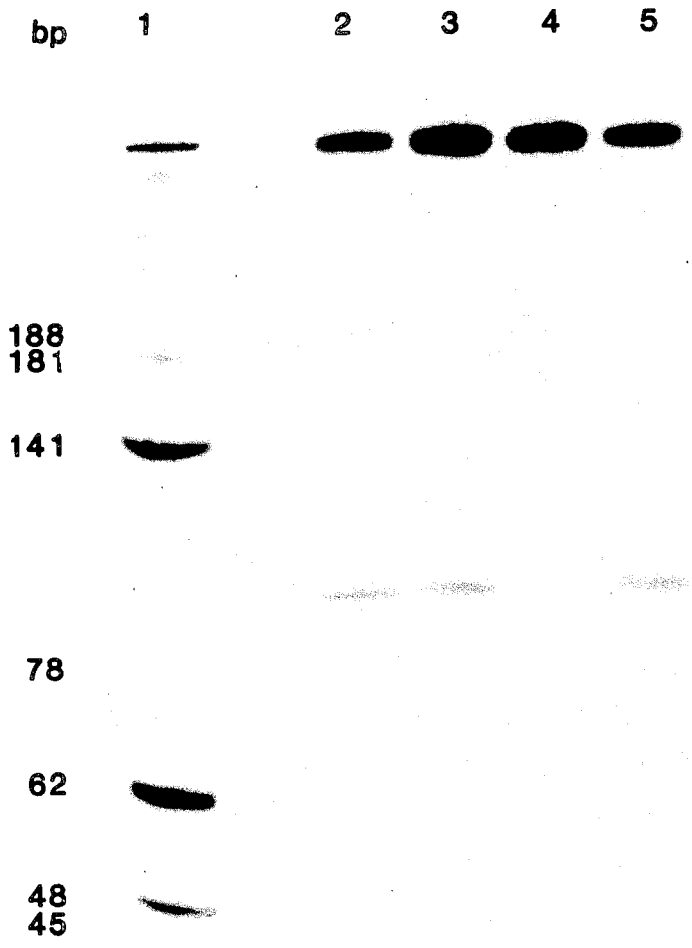


FIG.3

FIG.4-1

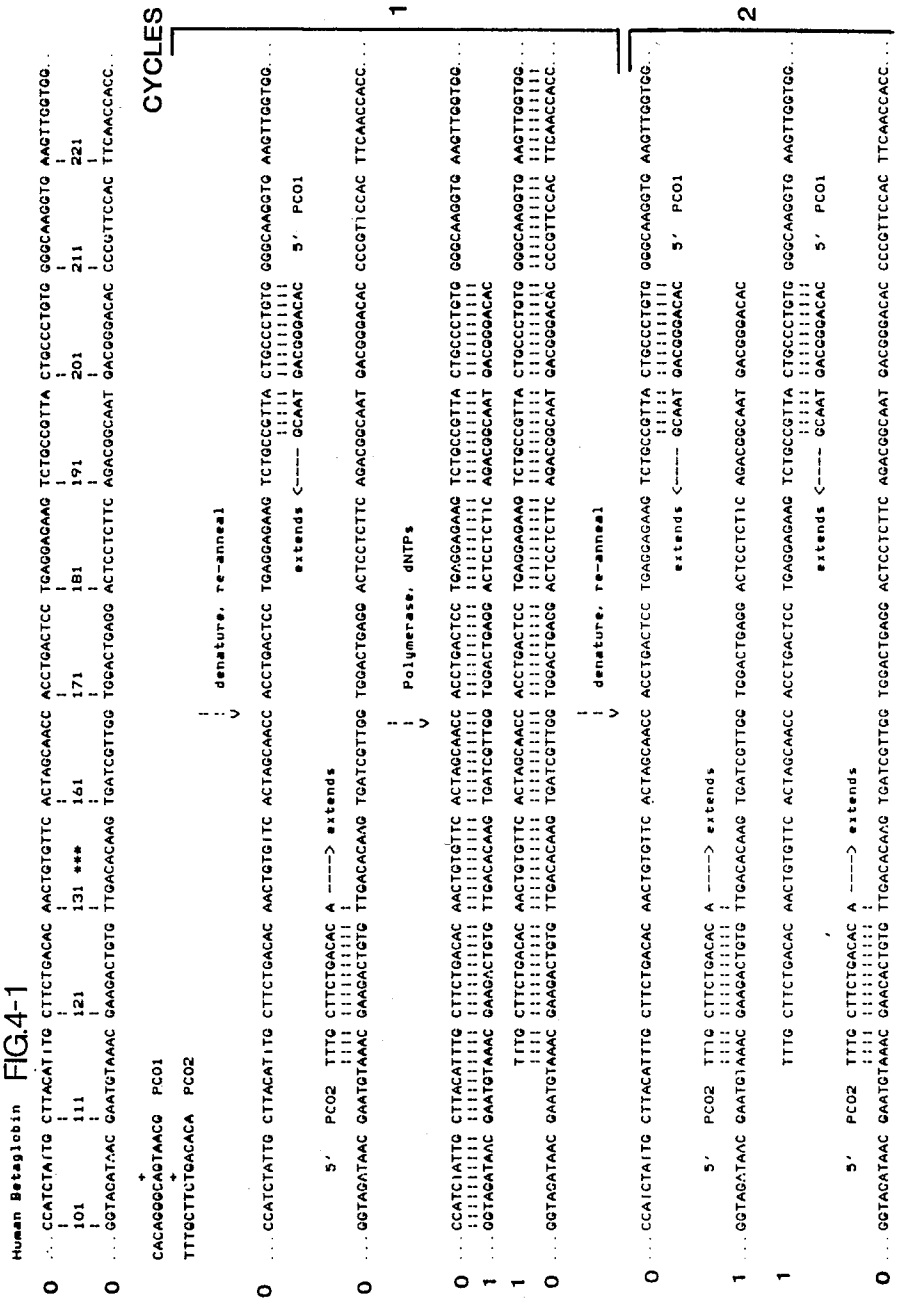

Human Beta-globin
CYCLES
PCO1
PCO2
denature, re-anneal
Polymerase, dNTPs
extends
5'
1
2

FIG.4-2

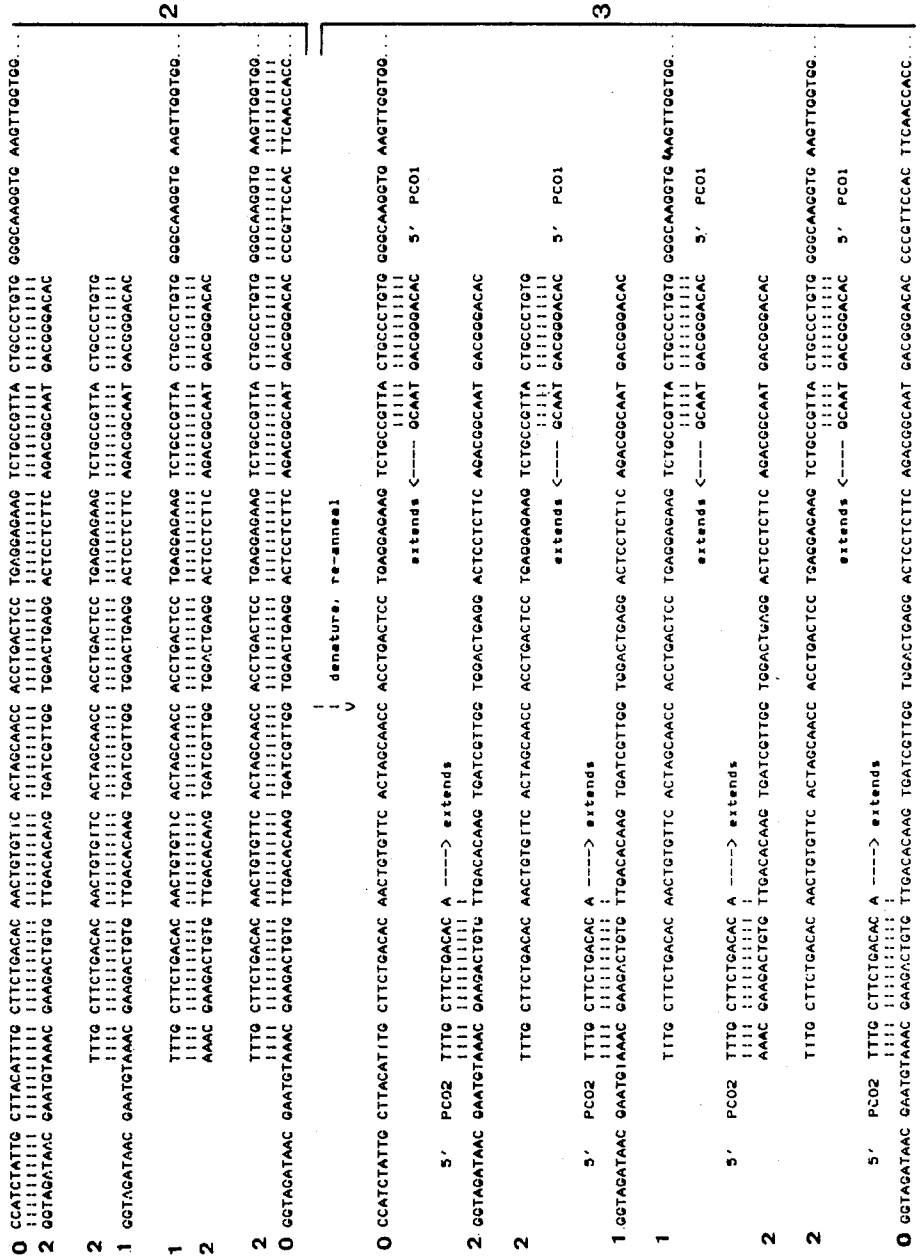
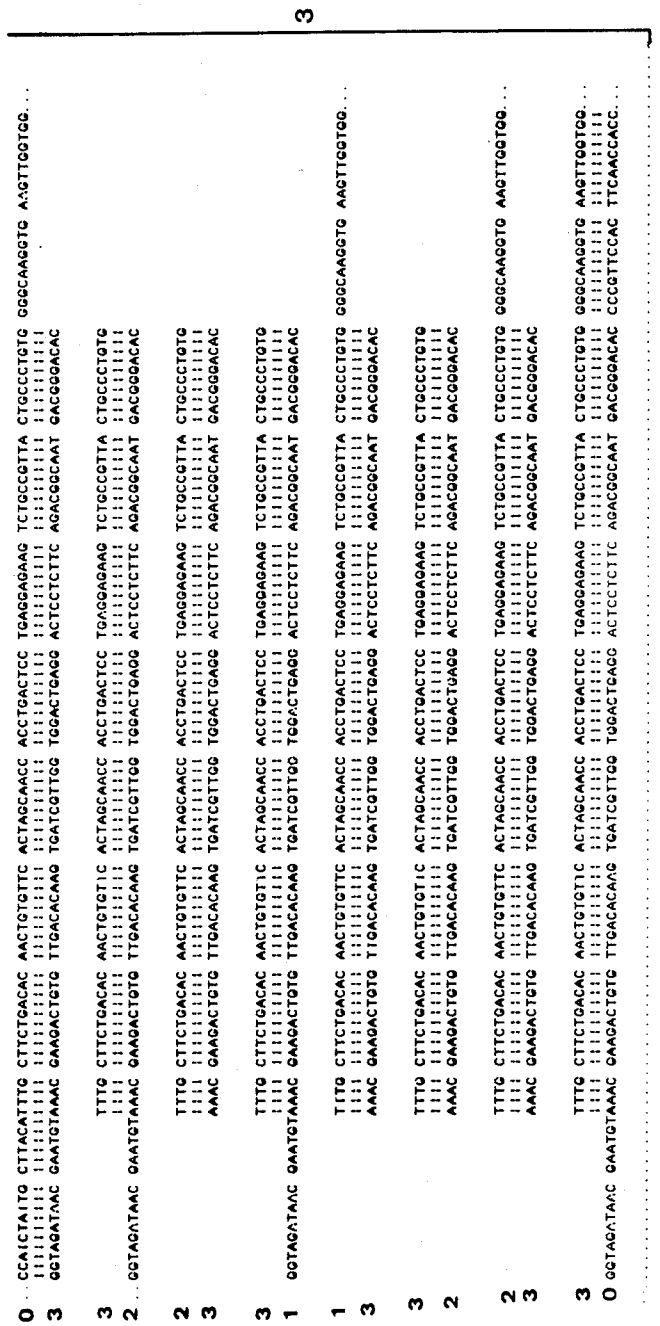


FIG.4-3


Polymerase, dNTPs
3
COPIES OF DNA SEQUENCE AFTER N CYCLES
N
0
1
5
10
15
20
template
long product
short product
=2(exp N)-N-1

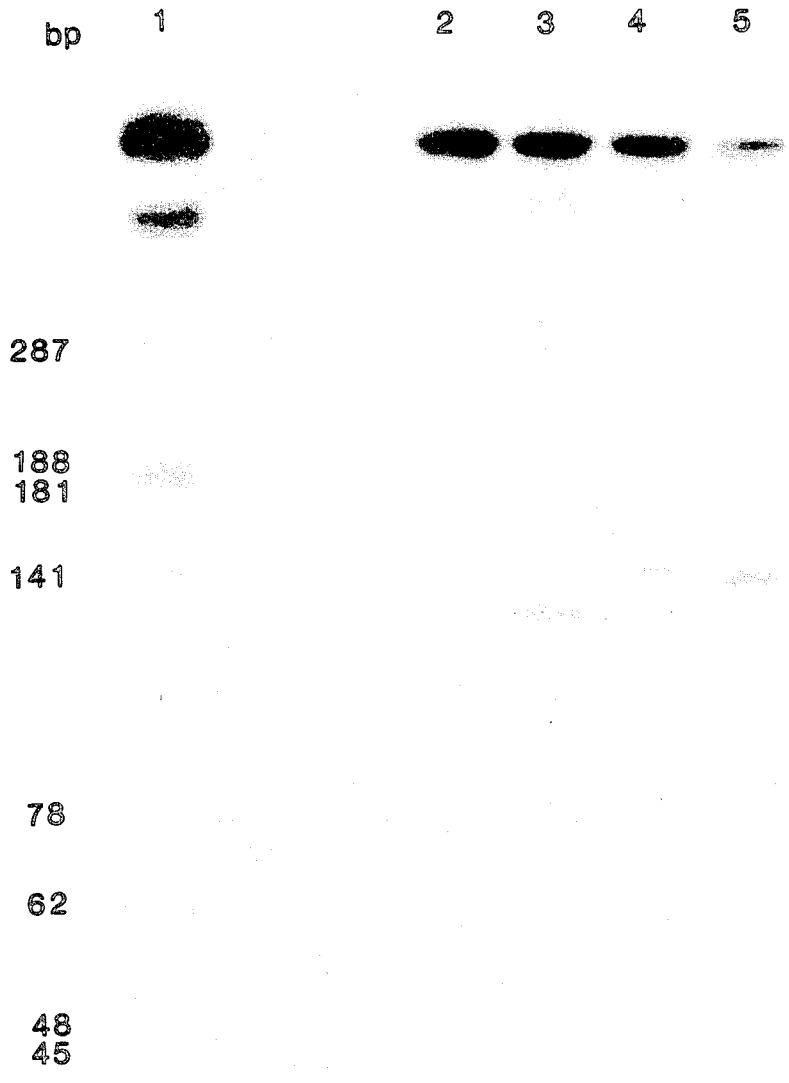


FIG.5

FIG.6

```
                    ==========
βA  CA TGG TGC ACC TGA C TCC TGA GGA GAA G TC TGC CG TTA C TGC CC TG TGG GGC AAG G TGA A
    G TAC CAC G TGG AC TGA GGA C TCC TC TTC AGA CGG CAA TGA CGG GAC ACC CCG TTC CAC TT
                    ==========

                    =====  *
βS  CA TGG TGC ACC TGA C TCC TG TGG AGA AG TC TGC CG TTA C TGC CC TG TGG GGC AAG G TGA A
    G TAC CAC G TGG AC TGA GGA C ACC TC TTC AGA CGG CAA TGA CGG GAC ACC CCG TTC CAC TT
                    =====
```

* Marks the mutation (A to T) in the sickle cell gene which disrupts the DdeI site

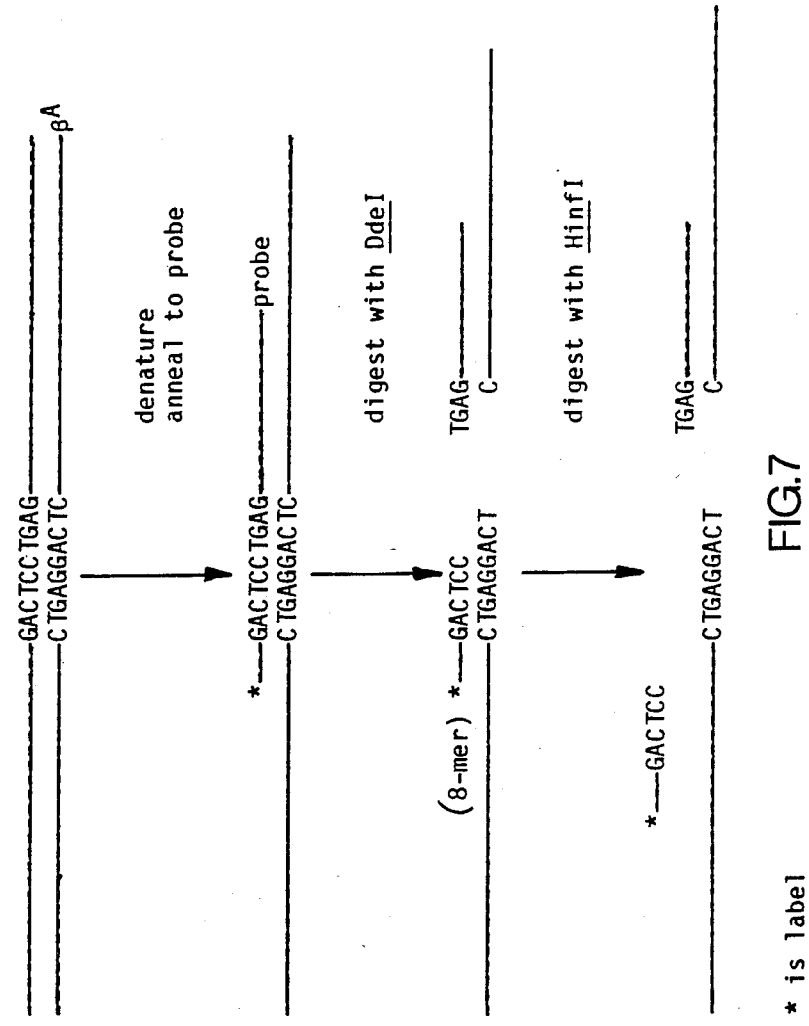

GACTCCTGAG
CTGAGGACTC
βA
denature
anneal to probe
*—GACTCCTGAG—probe
CTGAGGACTC
digest with DdeI
(8-mer) *—GACTCC
CTGAGGACT
TGAG—
C—
digest with HinfI
*—GACTCC
CTGAGGACT
TGAG—
C—
* is label

FIG.7

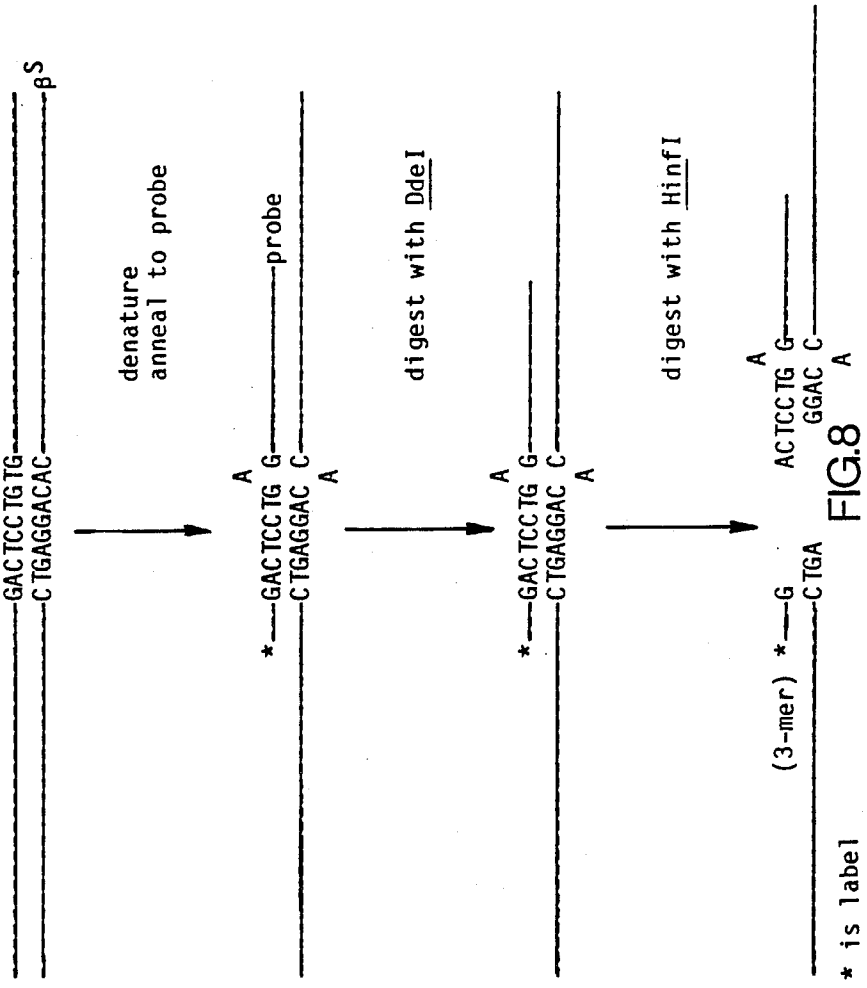

GACTCCTGTG
CTGAGGACAC
βS
denature
anneal to probe
probe
digest with DdeI
digest with HinfI
(3-mer)
ACTCCTG
GGAC
* is label

FIG.8

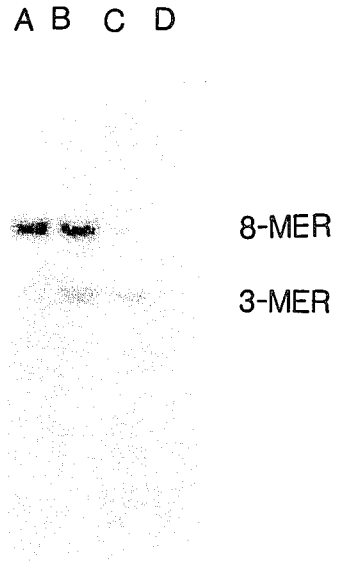
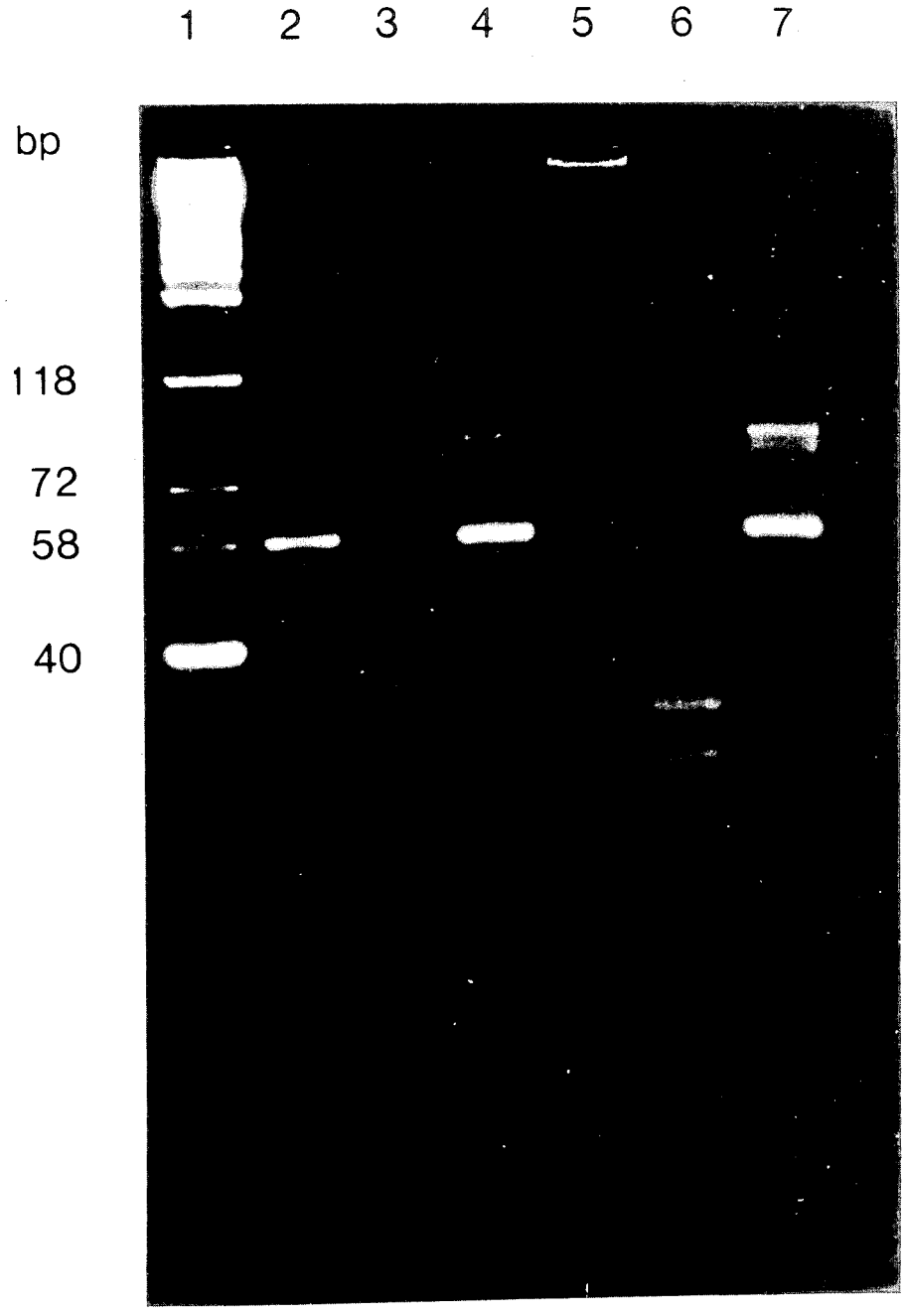


FIG.9

FIG. 10



PROCESS FOR AMPLIFYING NUCLEIC ACID SEQUENCES

This application is a continuation-in-part of copending U.S. application Ser. No. 716,975 filed Mar. 28, 1985, now abandoned.

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to a process for amplifying existing nucleic acid sequences. More specifically, it relates to a process for producing any particular nucleic acid sequence from a given sequence of DNA or RNA in amounts which are large compared to the amount initially present. The DNA or RNA may be single- or double-stranded, and may be a relatively pure species or a component of a mixture of nucleic acids. The process of the invention utilizes a repetitive reaction to accomplish the amplification of the desired nucleic acid sequence.

2. Description of Related Disclosures

For diagnostic applications in particular, the target nucleic acid sequence may be only a small portion of the DNA or RNA in question, so that it may be difficult to detect its presence using nonisotopically labeled or end-labeled oligonucleotide probes. Much effort is being expended in increasing the sensitivity of the probe detection systems, but little research has been conducted on amplifying the target sequence so that it is present in quantities sufficient to be readily detectable using currently available methods.

Several methods have been described in the literature for the synthesis of nucleic acids de novo or from an existing sequence. These methods are capable of producing large amounts of a given nucleic acid of completely specified sequence.

One known method for synthesizing nucleic acids de novo involves the organic synthesis of a nucleic acid from nucleoside derivatives. This synthesis may be performed in solution or on a solid support. One type of organic synthesis is the phosphotriester method, which has been utilized to prepare gene fragments or short genes. In the phosphotriester method, oligonucleotides are prepared which can then be joined together to form longer nucleic acids. For a description of this method, see Narang, S. A., et al., *Meth. Enzymol.*, 68, 90 (1979) and U.S. Pat. No. 4,356,270. The patent describes the synthesis and cloning of the somatostatin gene.

A second type of organic synthesis is the phosphodiester method, which has been utilized to prepare a tRNA gene. See Brown, E. L., et al., *Meth. Enzymol.*, 68, 109 (1979) for a description of this method. As in the phosphotriester method, the phosphodiester method involves synthesis of oligonucleotides which are subsequently joined together to form the desired nucleic acid.

Although the above processes for de novo synthesis may be utilized to synthesize long strands of nucleic acid, they are not very practical to use for the synthesis of large amounts of a nucleic acid. Both processes are laborious and time-consuming, require expensive equipment and reagents, and have a low overall efficiency. The low overall efficiency may be caused by the inefficiencies of the synthesis of the oligonucleotides and of the joining reactions. In the synthesis of a long nucleic acid, or even in the synthesis of a large amount of a shorter nucleic acid, many oligonucleotides would need

to be synthesized and many joining reactions would be required. Consequently, these methods would not be practical for synthesizing large amounts of any desired nucleic acid.

Methods also exist for producing nucleic acids in large amounts from small amounts of the initial existing nucleic acid. These methods involve the cloning of a nucleic acid in the appropriate host system, where the desired nucleic acid is inserted into an appropriate vector which is used to transform the host. When the host is cultured the vector is replicated, and hence more copies of the desired nucleic acid are produced. For a brief description of subcloning nucleic acid fragments, see Maniatis, T., et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, pp. 390–401 (1982). See also the techniques described in U.S. Pat. Nos. 4,416,988 and 4,403,036.

A third method for synthesizing nucleic acids, described in U.S. Pat. No. 4,293,652, is a hybrid of the above-described organic synthesis and molecular cloning methods. In this process, the appropriate number of oligonucleotides to make up the desired nucleic acid sequence is organically synthesized and inserted sequentially into a vector which is amplified by growth prior to each succeeding insertion.

The present invention bears some similarity to the molecular cloning method; however, it does not involve the propagation of any organism and thereby avoids the possible hazards or inconvenience which this entails. The present invention also does not require synthesis of nucleic acid sequences unrelated to the desired sequence, and thereby the present invention obviates the need for extensive purification of the product from a complicated biological mixture.

SUMMARY OF THE INVENTION

The present invention resides in a process for amplifying one or more specific nucleic acid sequences present in a nucleic acid or mixture thereof using primers and inducing agents. The extension product of one primer when hybridized to the other becomes a template for the production of the desired specific nucleic acid sequence, and vice versa, and the process is repeated as often as is necessary to produce the desired amount of the sequence. This method is expected to be more efficient than the methods described above for producing large amounts of nucleic acid from a target sequence and to produce such nucleic acid in a comparatively short period of time. The present method is especially useful for amplifying rare species of nucleic acid present in a mixture of nucleic acids for effective detection of such species.

More specifically, the present invention provides a process for amplifying at least one specific nucleic acid sequence contained in a nucleic acid or a mixture of nucleic acids wherein each nucleic acid consists of two separate complementary strands, of equal or unequal length, which process comprises:

(a) treating the strands with two primers, for each different specific sequence being amplified, under conditions such that for each different sequence being amplified an extension product of each primer is synthesized which is complementary to each nucleic acid strand, wherein said primers are selected so as to be substantially complementary to different strands of each specific sequence such that the extension product synthesized from one primer, when it is separated from its complement,

can serve as a template for synthesis of the extension product of the other primer;

(b) separating the primer extension products from the templates on which they were synthesized to produce single-stranded molecules; and

(c) treating the single-stranded molecules generated from step (b) with the primers of step (a) under conditions such that a primer extension product is synthesized using each of the single strands produced in step (b) as a template.

The steps may be conducted sequentially or simultaneously. In addition, steps (b) and (c) may be repeated until the desired level of sequence amplification is obtained.

In other embodiments the invention relates to methods for diagnosing the presence of specific nucleic acid sequences suspected of being in a sample and diagnostic kits applicable thereto.

The present invention may be useful not only for producing large amounts of an existing nucleic acid of completely specified sequence, but also for producing nucleic acid sequences which are known to exist but are not completely specified. In either case an initial copy of the sequence to be amplified must be available, although it need not be pure or a discrete molecule.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** illustrates a 94 base pair length sequence of human β-globin desired to be amplified. The single base pair change which is associated with sickle cell anemia is depicted beneath the 94-mer.

FIG. **2** illustrates an autoradiograph of polyacrylamide gel electrophoresis demonstrating amplification of the 94-mer contained in human wild-type DNA and in a plasmid containing a 1.9 kb BamHI fragment of the normal β-globin gene (pBR328:HbA).

FIG. **3** illustrates an autoradiograph of polyacrylamide gel electrophoresis demonstrating amplification of any of the specific target 94-mer sequence present in pBR328:HbA, a plasmid containing a 1.9 kb BamHI fragment of the sickle cell allele of β-globin (pBR328:HbS), pBR328:HbA where the sequence to be amplified is cleaved with MstII, and pBR328:HbS where the sequence to the amplified has been treated but not cleaved with MstII.

FIGS. **4-1-4-3** illustrate in detail the steps and products of the polymerase chain reaction for amplification of the desired 94-mer sequence of human β-globin for three cycles using two oligonucleotide primers.

FIG. **5** represents an autoradiograph of polyacrylamide gel electrophoresis demonstrating amplification after four cycles of a 240-mer sequence in pBR328:HbA, where the aliquots are digested with NcoI (Lane 3), MstII (Lane 4) or HinfI (Lane 5). Lane 1 is the molecular weight standard and Lane 2 contains the intact 240-bp product.

FIG. **6** illustrates the sequence of the normal (β^A) and sickle cell (β^S) β-globin genes in the region of the DdeI and HinfI restriction sites, where the single lines for β^A mark the position of the DdeI site (CTGAG) and the double bars for β^A and β^S mark the position of the HinfI site (GACTC).

FIG. **7** illustrates the results of sequential digestion of normal β-globin using a 40-mer probe and DdeI followed by HinfI restriction enzymes.

FIG. **8** illustrates the results of sequential digestion of sickle β-globin using the same 40-mer probe as in FIG. 7 and DdeI followed by HinfI restriction enzymes.

FIG. **9** illustrates an autoradiograph of polyacrylamide gel electrophoresis demonstrating the use of the same 40-mer probe as in FIG. 7 to specifically characterize the beta-globin alleles present in samples of whole human DNA which have been subjected to amplification by the present method.

FIG. **10** illustrates a photograph of a 6% NuSieve agarose gel visualized using ethidium bromide and UV light. This photograph demonstrates amplification of a sub-fragment of a 110-bp amplification product which sub-fragment is an inner nested set within the 110-bp fragment.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The term "oligonucleotide" as used herein in referring to primers, probes, oligomer fragments to be detected, oligomer controls and unlabeled blocking oligomers is defined as a molecule comprised of two or more deoxyribonucleotides or ribonucleotides, preferably more than three. Its exact size will depend on many factors, which in turn depend on the ultimate function or use of the oligonucleotide.

The term "primer" as used herein refers to an oligonucleotide, whether occurring naturally as in a purified restriction digest or produced synthetically, which is capable of acting as a point of initiation of synthesis when placed under conditions in which synthesis of a primer extension product which is complementary to a nucleic acid strand is induced, i.e., in the presence of nucleotides and an inducing agent such as DNA polymerase and at a suitable temperature and pH. The primer is preferably single stranded for maximum efficiency in amplification, but may alternatively be double stranded. If double stranded, the primer is firt treated to separate its strands before being used to prepare extension products. Preferably, the primer is an oligodeoxyribonucleotide. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the inducing agent. The exact lengths of the primers will depend on many factors, including temperature, source of primer and use of the method. For example, for diagnostics applications, depending on the complexity of the target sequence, the oligonucleotide primer typically contains 15–25 or more nucleotides, although it may contain fewer nucleotides. For other applications, the oligonucleotide primer is typically shorter, e.g., 7–15 nucleotides. Such short primer molecules generally require cooler temperatures to form sufficiently stable hybrid complexes with template.

The primers herein are selected to be "substantially" complementary to the different strands of each specific sequence to be amplified. This means that the primers must be sufficiently complementary to hybridize with their respective strands. Therefore, the primer sequence need not reflect the exact sequence of the template. For example, a non-complementary nucleotide fragment may be attached to the 5' end of the primer, with the remainder of the primer sequence being complementary to the strand. Alternatively, non-complementary bases or longer sequences can be interspersed into the primer, provided that the primer sequence has sufficient complementarity with the sequence of the strand to be amplified to hybridize therewith and thereby form a template for synthesis of the extension product of the other primer.

As used herein, the terms "restriction endonucleases" and "restriction enzymes" refer to bacterial enzymes

each of which cut double-stranded DNA at or near a specific nucleotide sequence.

As used herein, the term "DNA polymorphism" refers to the condition in which two or more different nucleotide sequences coexists in the same interbreeding population in a DNA sequence.

The term "restriction fragment length polymorphism" ("RFLP") refers to the differences in DNA nucleotide sequences that are randomly distributed throughout the entire human genome and that produce different restriction endonuclease patterns.

The present invention is directed to a process for amplifying any one or more desired specific nucleic acid sequences found in a nucleic acid. Because large amounts of a specific sequence may be produced by this process, the present invention may be used for improving the efficiency of cloning DNA or messenger RNA and for amplifying a target sequence to facilitate detection thereof. The present invention is also useful for obtaining large amounts of the desired sequence from a mixture of nucleic acids resulting from an imperfect chemical synthesis.

In general, the present process involves a chain reaction for producing, in exponential quantities relative to the number of reaction steps involved, at least one specific nucleic acid sequence given (a) that the ends of the required sequence are known in sufficient detail that oligonucleotides can be synthesized which will hybridize to them, and (b) that a small amount of the sequence is available to initiate the chain reaction. The product of the chain reaction will be a discrete nucleic acid duplex with termini corresponding to the ends of the specific primers employed.

Any source of nucleic acid, in purified or nonpurified form, can be utilized as the starting nucleic acid or acids, provided it contains or is suspected of containing the specific nucleic acid sequence desired. Thus, the process may employ, for example, DNA or RNA, including messenger RNA, which DNA or RNA may be single stranded or double stranded. In addition, a DNA-RNA hybrid which contains one strand of each may be utilized. A mixture of any of these nucleic acids may also be employed, or the nucleic acid produced from a previous amplification reaction herein using the same or different primers may be so utilized. The specific nucleic acid sequence to be amplified may be only a fraction of a larger molecule or can be present initially as a discrete molecule, so that the specific sequence constitutes the entire nucleic acid. It is not necessary that the sequence to be amplified be present initially in a pure form; it may be a minor fraction of a complex mixture, such as a portion of the β-globin gene contained in whole human DNA or a portion of nucleic acid sequence due to a particular microorganism which organism might constitute only a very minor fraction of a particular biological sample. The starting nucleic acid may contain more than one desired specific nucleic acid sequence which may be the same or different. Therefore, the present process is useful not only for producing large amounts of one specific nucleic acid sequence, but also for amplifying simultaneously more than one different specific nucleic acid sequence located on the same or different nucleic acid molecules.

The nucleic acid or acids may be obtained from any source, for example, from plasmids such as pBR322, from cloned DNA or RNA, or from natural DNA or RNA from any source, including bacteria, yeast, viruses, and higher organisms such as plants or animals. DNA or RNA may be extracted from blood, tissue material such as chorionic villi or amniotic cells by a variety of techniques such as that described by Maniatis et al., *Molecular Cloning A Laboratory Manual* (New York: Cold Spring Harbor Laboratory, 1982), pp. 280–281.

Any specific nucleic acid sequence can be produced by the present process. It is only necessary that a sufficient number of bases at both ends of the sequence be known in sufficient detail so that two oligonucleotide primers can be prepared which will hybridize to different strands of the desired sequence and at relative positions along the sequence such that an extension product synthesized from one primer, when it is separated from its template (complement), can serve as a template for extension of the other primer into a nucleic acid of defined length. The greater the knowledge about the bases at both ends of the sequence, the greater can be the specificity of the primers for the target nucleic acid sequence, and thus the greater the efficiency of the process. It will be understood that the word primer as used hereinafter may refer to more than one primer, particularly in the case where there is some ambiguity in the information regarding the terminal sequence(s) of the fragment to be amplified. For instance, in the case where a nucleic acid sequence is inferred from protein sequence information a collection of primers containing sequences representing all possible codon variations based on degeneracy of the genetic code will be used for each strand. One primer from this collection will be 100% homologous with the end of the desired sequence to be amplified.

The oligonucleotide primers may be prepared using any suitable method, such as, for example, the phosphotriester and phosphodiester methods described above, or automated embodiments thereof. In one such automated embodiment diethylphosphoramidites are used as starting materials and may be synthesized as described by Beaucage et al., *Tetrahedron Letters* (1981), 22: 1859–1962. One method for synthesizing oligonucleotides on a modified solid support is described in U.S. Pat. No. 4,458,066. It is also possible to use a primer which has been isolated from a biological source (such as a restriction endonuclease digest).

The specific nucleic acid sequence is produced by using the nucleic acid containing that sequence as a template. If the nucleic acid contains two strands, it is necessary to separate the strands of the nucleic acid before it can be used as the template, either as a separate step or simultaneously with the synthesis of the primer extension products. This strand separation can be accomplished by any suitable method including physical, chemical or enzymatic means. One physical method of separating the strands of the nucleic acid involves heating the nucleic acid until it is completely (>99%) denatured. Typical heat denaturation may involve temperatures ranging from about 80° to 105° C. for times ranging from about 1 to 10 minutes. Strand separation may also be induced by an enzyme from the class of enzymes known as helicases or the enzyme RecA, which has helicase activity and in the presence of riboATP is known to denature DNA. The reaction conditions suitable for separating the strands of nucleic acids with helicases are described by Cold Spring Harbor Symposia on Quantitative Biology, Vol. XLIII "DNA: Replication and Recombination" (New York: Cold Spring Harbor Laboratory, 1978), B. Kuhn et al., "DNA Helicases", pp. 63–67, and techniques for using RecA are

reviewed in C. Radding, *Ann. Rev. Genetics.* 16: 405–37 (1982).

If the original nucleic acid containing the sequence to be amplified is single stranded, its complement is synthesized by adding one or two oligonucleotide primers thereto. If an appropriate single primer is added, a primer extension product is synthesized in the presence of the primer, an inducer or catalyst of the synthesis and the four nucleotides described below. The product will be partially complementary to the single-stranded nucleic acid and will hybridize with the nucleic acid strand to form a duplex of unequal length strands that may then be separated into single strands as described above to produce two single separated complementary strands. Alternatively, two appropriate primers may be added to the single-stranded nucleic acid and the reaction carried out.

If the original nucleic acid constitutes the sequence to be amplified, the primer extension product(s) produced will be completely complementary to the strands of the original nucleic acid and will hybridize therewith to form a duplex of equal length strands to be separated into single-stranded molecules.

When the complementary strands of the nucleic acid or acids are separated, whether the nucleic acid was originally double or single stranded, the strands are ready to be used as a template for the synthesis of additional nucleic acid strands. This synthesis can be performed using any suitable method. Generally it occurs in a buffered aqueous solution, preferably at a pH of 7–9, most preferably about 8. Preferably, a molar excess (for cloned nucleic acid, usually about 1000:1 primer:template, and for genomic nucleic acid, usually about 10^6:1 primer:template) of the two oligonucleotide primers is added to the buffer containing the separated template strands. It is understood, however, that the amount of complementary strand may not be known if the process herein is used for diagnostic applications, so that the amount of primer relative to the amount of complementary strand cannot be determined with certainty. As a practical matter, however, the amount of primer added will generally be in molar excess over the amount of complementary strand (template) when the sequence to be amplified is contained in a mixture of complicated long-chain nucleic acid strands. A large molar excess is preferred to improve the efficiency of the process.

The deoxyribonucleoside triphosphates dATP, dCTP, dGTP and TTP are also added to the synthesis mixture in adequate amounts and the resulting solution is heated to about 90°–100° C. for from about 1 to 10 minutes, preferably from 1 to 4 minutes. After this heating period the solution is allowed to cool to room temperature, which is preferable for the primer hybridization. To the cooled mixture is added an appropriate agent for inducing or catalyzing the primer extension reaction, and the reaction is allowed to occur under conditions known in the art. This synthesis reaction may occur at from room temperature up to a temperature above which the inducing agent no longer functions efficiently. Thus, for example, if DNA polymerase is used as inducing agent, the temperature is generally no greater than about 40° C. Most conveniently the reaction occurs at room temperature.

The inducing agent may be any compound or system which will function to accomplish the synthesis of primer extension products, including enzymes. Suitable enzymes for this purpose include, for example, *E. coli* DNA polymerase I, Klenow fragment of *E. coli* DNA polymerase I, T4 DNA polymerase, other available DNA polymerases, reverse transcriptase, and other enzymes, including heat-stable enzymes, which will facilitate combination of the nucleotides in the proper manner to form the primer extension products which are complementary to each nucleic acid strand. Generally, the synthesis will be initiated at the 3′ end of each primer and proceed in the 5′ direction along the template strand, until synthesis terminates, producing molecules of different lengths. There may be inducing agents, however, which initiate synthesis at the 5′ end and proceed in the other direction, using the same process as described above.

The newly synthesized strand and its complementary nucleic acid strand form a double-stranded molecule which is used in the succeeding steps of the process. In the next step, the strands of the double-stranded molecule are separated using any of the procedures described above to provide single-stranded molecules.

New nucleic acid is synthesized on the single-stranded molecules. Additional inducing agent, nucleotides and primers may be added if necessary for the reaction to proceed under the conditions prescribed above. Again, the synthesis will be initiated at one end of the oligonucleotide primers and will proceed along the single strands of the template to produce additional nucleic acid. After this step, half of the extension product will consist of the specific nucleic acid sequence bounded by the two primers.

The steps of strand separation and extension product synthesis can be repeated as often as needed to produce the desired quantity of the specific nucleic acid sequence. As will be described in further detail below, the amount of the specific nucleic acid sequence produced will accumulate in an exponential fashion.

When it is desired to produce more than one specific nucleic acid sequence from the first nucleic acid or mixture of nucleic acids, the appropriate number of different oligonucleotide primers are utilized. For example, if two different specific nucleic acid sequences are to be produced, four primers are utilized. Two of the primers are specific for one of the specific nucleic acid sequences and the other two primers are specific for the second specific nucleic acid sequence. In this manner, each of the two different specific sequences can be produced exponentially by the present process.

The present invention can be performed in a step-wise fashion where after each step new reagents are added, or simultaneously, where all reagents are added at the initial step, or partially step-wise and partially simultaneous, where fresh reagent is added after a given number of steps. If a method of strand separation, such as heat, is employed which will inactivate the inducing agent, as in the case of a heat-labile enzyme, then it is necessary to replenish the inducing agent after every strand separation step. The simultaneous method may be utilized when an enzymatic means is used for the strand separation step. In the simultaneous procedure, the reaction mixture may contain, in addition to the nucleic acid strand(s) containing the desired sequence, the strand-separating enzyme (e.g., helicase), an appropriate energy source for the strand-separating enzyme, such as rATP, the four nucleotides, the oligonucleotide primers in molar excess, and the inducing agent, e.g., Klenow fragment of *E. coli* DNA polymerase I. If heat is used for denaturation in a simultaneous process, a heat-stable inducing agent such as a thermostable poly-

merase may be employed which will operate at an elevated temperature, preferably 65°–90° C. depending on the inducing agent, at which temperature the nucleic acid will consist of single and double strands in equilibrium. For smaller lengths of nucleic acid, lower temperatures of about 50° C. may be employed. The upper temperature will depend on the temperature at which the enzyme will degrade or the temperature above which an insufficient level of primer hybridization will occur. Such a heat-stable enzyme is described, e.g., by A. S. Kaledin et al., *Biokhimiya,* 45, 644–651 (1980). Each step of the process will occur sequentially notwithstanding the initial presence of all the reagents. Additional materials may be added as necessary. After the appropriate length of time has passed to produce the desired amount of the specific nucleic acid sequence, the reaction may be halted by inactivating the enzymes in any known manner or separating the components of the reaction.

The process of the present invention may be conducted continuously. In one embodiment of an automated process, the reaction may be cycled through a denaturing region, a reagent addition region, and a reaction region. In another embodiment, the enzyme used for the synthesis of primer extension products can be immobilized in a column. The other reaction components can be continuously circulated by a pump through the column and a heating coil in series; thus the nucleic acids produced can be repeatedly denatured without inactivating the enzyme.

desired sequence [S] comprised of complementary strands [S+] and [S−] is utilized as the nucleic acid. During the first and each subsequent reaction cycle extension of each oligonucleotide primer on the original template will produce one new ssDNA molecule product of indefinite length which terminates with only one of the primers. These products, hereafter referred to as "long products," will accumulate in a linear fashion; that is, the amount present after any number of cycles will be proportional to the number of cycles.

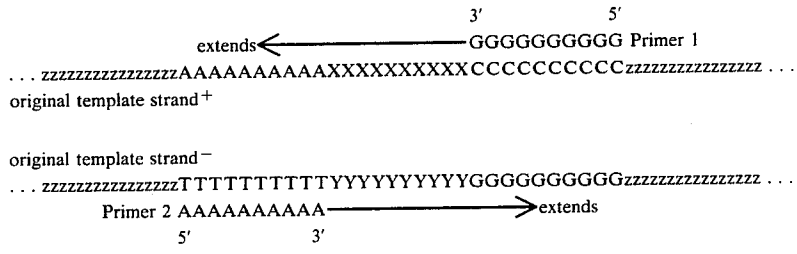
The long products thus produced will act as templates for one or the other of the oligonucleotide primers during subsequent cycles and will produce molecules of the desired sequence [S+] or [S−]. These molecules will also function as templates for one or the other of the oligonucleotide primers, producing further [S+] and [S−], and thus a chain reaction can be sustained which will result in the accumulation of [S] at an exponential rate relative to the number of cycles.

By-products formed by oligonucleotide hybridizations other than those intended are not self-catalytic (except in rare instances) and thus accumulate at a linear rate.

The specific sequence to be amplified, [S], can be depicted diagrammatically as:

[S+] 5′ AAAAAAAAAAXXXXXXXXXXCCCCCCCCCC 3′
[S−] 3′ TTTTTTTTTTYYYYYYYYYYGGGGGGGGGG 5′
The appropriate oligonucleotide primers would be:
Primer 1: GGGGGGGGGG
Primer 2: AAAAAAAAAA
so that if DNA containing [S]
. . . zzzzzzzzzzzzzzzzAAAAAAAAAAXXXXXXXXXXCCCCCCCCCCzzzzzzzzzzzzzzzz . . .
. . . zzzzzzzzzzzzzzzzTTTTTTTTTTYYYYYYYYYYGGGGGGGGGGzzzzzzzzzzzzzzzz . . .

The present invention is demonstrated diagrammatically below were double-stranded DNA containing the desired sequence [S] is separated into single strands and its single strands are hybridized to Primers 1 and 2, the following extension reactions can be catalyzed by DNA polymerase in the presence of the four deoxyribonucleoside triphosphates:



On denaturation of the two duplexes formed, the products are:

3′ 5′
. . .zzzzzzzzzzzzzzzzTTTTTTTTTTYYYYYYYYYYGGGGGGGGGG
newly synthesized long product 1

5′ 3′
. . .zzzzzzzzzzzzzzzzAAAAAAAAAAXXXXXXXXXXCCCCCCCCCCzzzzzzzzzzzzzzzz . . .
original template strand+

Case 3:11-cv-00056-IEG -WVG Document 1-2 Filed 01/10/11 Page 77 of 90

-continued

3′ 5′
. . .zzzzzzzzzzzzzzzzTTTTTTTTTTTYYYYYYYYYYGGGGGGGGGGzzzzzzzzzzzzzzzz . . .
original template strand

5′ 3′
AAAAAAAAAAXXXXXXXXXXCCCCCCCCCCzzzzzzzzzzzzzzzz . . .
newly synthesized long product 3

If these four strands are allowed to rehybridize with Primers 1 and 2 in the next cycle, inducing agent will catalyze the following reactions:

Primer 2 5′ AAAAAAAAAA ——————→ extends to here
3′ . . . zzzzzzzzzzzzzzzzzzTTTTTTTTTTTYYYYYYYYYYGGGGGGGGGG 5′
newly synthesized long product 1

extends ←—————— GGGGGGGGGG 5′ Primer 1
5′ . . .zzzzzzzzzzzzzzAAAAAAAAAAXXXXXXXXXXCCCCCCCCCCzzzzzzzzzzzzzz . . . 3′
original template strand+

Primer 2 5′ AAAAAAAAAA ——————→ extends
3′ . . . zzzzzzzzzzzzzzzzzzTTTTTTTTTTTYYYYYYYYYYGGGGGGGGGGzzzzzzzzzz . . . 5′
original template strand−

extends to here ←—————— GGGGGGGGGG 5′ Primer 1
5′ AAAAAAAAAAXXXXXXXXXXCCCCCCCCCCzzzzzzzzzzzzzzzz . . . 3′
newly synthesized long product 2

If the strands of the above four duplexes are separated, the following strands are found:

5′ AAAAAAAAAAXXXXXXXXXXCCCCCCCCCC 3′
newly synthesized [S+]
3′ . . . zzzzzzzzzzzzzzzzzzTTTTTTTTTTYYYYYYYYYYGGGGGGGGGG 5′
first cycle synthesized long product 1
3′ . . . zzzzzzzzzzzzzzzzzzTTTTTTTTTTYYYYYYYYYYGGGGGGGGGG 5′
newly synthesized long product 1
5′ . . . zzzzzzzzzzzzzzzzzzzzAAAAAAAAAAXXXXXXXXXXCCCCCCCCCCzzzzzzzzz . . . 3′
original template strand+
5′ AAAAAAAAAAXXXXXXXXXXCCCCCCCCCCzzzzzzzzzzzzzzzz . . . 3′
newly synthesized long product 2
3′ . . . zzzzzzzzzzzzzzzTTTTTTTTTTYYYYYYYYYYGGGGGGGGGGzzzzzzzzzzzzzzzz . . . 5′
original template strand−
3′ TTTTTTTTTTYYYYYYYYYYGGGGGGGGGG 5′
newly synthesized [S−]
5′ AAAAAAAAAAXXXXXXXXXXCCCCCCCCCCzzzzzzzzzzzzzzzz . . . 3′
first cycle synthesized long product 2

Cycle Number	Template	Long Products	Specific Sequence [S]
0	1	—	—
1	1	1	0
2	1	2	1
3	1	3	4
5	1	5	26
10	1	10	1013
15	1	15	32,752
20	1	20	1,048,555
n	1	n	(2^n-n-1)

It is seen that each strand which terminates with the oligonucleotide sequence of one primer and the complementary sequence of the other is the specific nucleic acid sequence [S] that is desired to be produced.

The steps of this process can be repeated indefinitely, being limited only by the amount of Primers 1 and 2, inducing agent and nucleotides present. The amount of original nucleic acid remains constant in the entire process, because it is not replicated. The amount of the long products increases linearly because they are produced only from the original nucleic acid. The amount of the specific sequence increases exponentially. Thus, the specific sequence will become the predominant species. This is illustrated in the following table, which indicates the relative amounts of the species theoretically present after n cycles, assuming 100% efficiency at each cycle:

When a single-stranded nucleic acid is utilized as the template, only one long product is formed per cycle.

The method herein may be utilized to clone a particular nucleic acid sequence for insertion into a suitable expression vector. The vector may then be used to transform an appropriate host organism to produce the gene product of the sequence by standard methods of recombinant DNA technology.

In addition, the process herein can be used for in vitro mutagenesis. The oligodeoxyribonucleotide primers need not be exactly complementary to the DNA se-

quence which is being amplified. It is only necessary that they be able to hydridize to the sequence sufficiently well to be extended by the polymerase enzyme or by whatever other inducing agent is employed. The product of a polymerase chain reaction wherein the primers employed are not exactly complementary to the original template will contain the sequence of the primer rather than the template, thereby introducing an in vitro mutation. In further cycles this mutation will be amplified with an undiminished efficiency because no further mispaired primings are required. The mutant thus produced may be inserted into an appropriate vector by standard molecular biological techniques and might confer mutant properties on this vector such as the potential for production of an altered protein.

The process of making an altered DNA sequence as described above could be repeated on the altered DNA using different primers so as to induce further sequence changes. In this way a series of mutated sequences could gradually be produced wherein each new addition to the series could differ from the last in a minor way, but from the original DNA source sequence in an increasingly major way. In this manner changes could be made ultimately which were not feasible in a single step due to the inability of a very seriously mismatched primer to function.

In addition, the primer can contain as part of its sequence a non-complementary sequence provided that a sufficient amount of the primer contains a sequence which is complementary to the strand to be amplified. For example, a nucleotide sequence which is not complementary to the template sequence (such as, e.g., a promoter, linker, coding sequence, etc.) may be attached at the 5' end of one or both of the primers, and thereby appended to the product of the amplification process. After the extension primer is added, sufficient cycles are run to achieve the desired amount of new template containing the non-complementary nucleotide insert. This allows production of large quantities of the combined fragments in a relatively short period of time (e.g., two hours or less) using a simple technique.

The method herein may also be used to enable detection and/or characterization of specific nucleic acid sequences associated with infectious diseases, genetic disorders or cellular disorders such as cancer. Amplification is useful when the amount of nucleic acid available for analysis is very small, as, for example, in the prenatal diagnosis of sickle cell anemia using DNA obtained from fetal cells. Amplification is particularly useful if such an analysis is to be done on a small sample using non-radioactive detection techniques which may be inherently insensitive, or where radioactive techniques are being employed but where rapid detection is desirable.

For purposes of this invention genetic diseases may include specific deletions and/or mutations in genomic DNA from any organism, such as, e.g., sickle cell anemia, cystic fibrosis, α-thalassemia, β-thalassemia, and the like. Sickle cell anemia can be readily detected via oligomer restriction analysis or a RFLP-like analysis following amplification of the appropriate DNA sequence by the present method. α-Thalassemia can be detected by the absence of a sequence, and β-thalassemia can be detected by the presence of a polymorphic restriction site closely linked to a mutation which causes the disease.

All of these genetic diseases may be detected by amplifying the appropriate sequence and analyzing it by Southern blots without using radioactive probes. In such a process, for example, a small sample of DNA from, e.g., amniotic fluid containing a very low level of the desired sequence is amplified, cut with a restriction enzyme, and analyzed via a Southern blotting technique. The use of non-radioactive probes is facilitated by the high level of the amplified signal.

In another embodiment a small sample of DNA may be amplified to a convenient level and then a further cycle of extension reactions performed wherein nucleotide derivatives which are readily detectable (such as ^{32}P-labeled or biotin labeled nucleoside triphosphates) are incorporated directly into the final DNA product, which may be analyzed by restriction and electrophoretic separation or any other appropriate method. An example of this technique in a model system is demonstrated in FIG. **5**.

In a further embodiment, demonstrated in a model system in FIG. **3**, the nucleic acid may be exposed to a particular restriction endonuclease prior to amplification. Since a sequence which has been cut cannot be amplified, the appearance of an amplified fragment, despite prior restriction of the DNA sample, implies the absence of a site for the endonuclease within the amplified sequence. The presence or absence of an amplified sequence can be detected by an appropriate method.

A practical application of this technique can be illustrated by its use in facilitating the detection of sickle cell anemia via the oligomer restriction technique described herein and in copending U.S. application Ser. No. 716,982 to Erlich et al. entitled "Method For Detection of Polymorphic Restriction Sites and Nucleic Acid Sequences" filed Mar. 28, 1985. Sickle cell anemia is a hemoglobin disease which is caused by a single base pair change in the sixth codon of the β-globin gene. FIG. **6** illustrates the sequences of normal and sickle cell β-globin genes in the region of their polymorphism, where the single bars mark the location of a DdeI site present only in the normal gene and where the double bars mark the location of a HinfI site which is non-polymorphic and thus present in both the normal and sickle cell alleles. FIG. **7** illustrates the process of oligomer restriction of normal β-globin DNA using a probe spanning both restriction sites and labeled where the asterisk appears. The DNA, amplified as provided herein, is denatured and annealed to the labeled probe. The enzyme DdeI cleaves the DNA at the reformed DdeI site and generates a labeled octamer. Under the conditions used in the test the octamer is short enough to dissociate from the duplex. The subsequent addition of the enzyme HinfI has no effect on the now single-stranded octamer. FIG. **8** illustrates the same process applied to the sickle cell allele of β-globin DNA. The enzyme DdeI cannot cleave the duplex formed by the amplified DNA and the labeled probe because of the A—A base pair mismatch. The enzyme HinfI, however, does restrict the hybrid and a labeled trimer is produced. In practice the method can diagnose the DNA of an individual as being either homozygous for the wild type, homozygous for the sickle type or a heterozygous carrier of the sickle cell trait, since a specific signal is associated with the presence of either allele. Use of this above-described method to amplify the pertinent sequence allows for a rapid analysis of a single copy gene using a probe with only a single ^{32}P label.

Various infectious diseases can be diagnosed by the presence in clinical samples of specific DNA sequences characteristic of the causative microorganism. These

include bacteria, such as Salmonella, Chlamydia, and Neisseria; viruses, such as the hepatitis viruses; and protozoan parasites, such as the Plasmodium responsible for malaria. U.S. Pat. No. 4,358,535 issued to Falkow describes the use of specific DNA hybridization probes for the diagnosis of infectious diseases. A problem inherent in the Falkow procedure is that a relatively small number of pathogenic organisms may be present in a clinical sample from an infected patient and the DNA extracted from these may constitute only a very small fraction of the total DNA in the sample. Specific amplification of suspected sequences prior to immobilization and hybridization detection of the DNA samples could greatly improve the sensitivity and specificity of these procedures.

Routine clinical use of DNA probes for the diagnosis of infectious diseases would be simplified considerably if non-radioactively labeled probes could be employed as described in EP 63,879 to Ward. In this procedure biotin-containing DNA probes are detected by chromogenic enzymes linked to avidin or biotin-specific antibodies. This type of detection is convenient, but relatively insensitive. The combination of specific DNA amplification by the present method and the use of stably labeled probes could provide the convenience and sensitivity required to make the Falkow and Ward procedures useful in a routine clinical setting.

The amplification process can also be utilized to produce sufficient quantities of DNA from a single copy human gene such that detection by a simple non-specific DNA stain such as ethidium bromide can be employed so as to make a DNA diagnosis directly.

In addition to detecting infectious diseases and pathological abnormalities in the genome of organisms, the process herein can also be used to detect DNA polymorphism which may not be associated with any pathological state.

The following examples are offered by way of illustration and are not intended in limit the invention in any manner. In these examples all percentages are by weight if for solids and by volume if for liquids, and all temperatures are in degrees Celsius unless otherwise noted.

EXAMPLE 1

A 25 base pair sequence having the nucleotide sequence

5′ CCTCGGCACCGTCACCCTGGATGCT 3′

3′ GGAGCCGTGGCAGTGGGACCTACGA 5′

contained on a 47 base pair FokI restriction fragment of pBR322 obtainable from ATCC was prepared as follows. A FokI digest of pBR322 containing the 47-bp fragment was produced by digesting pBR322 with FokI in accordance with the conditions suggested by the supplier, New England Biolabs Inc. The primers which were utilized were 5′ d(CCTCGGCACCG) 3′ and 5′ d(AGCATCCAGGGTG) 3′, and were prepared using conventional techniques. The following ingredients were added to 33 μl of buffer which consisted of 25 mM potassium phosphate, 10 mM magnesium chloride and 100 mM sodium chloride at pH 7.5: 2433 pmoles of each of the primers described above, 2.4 pmoles of the FokI digest of pBR322, 12 nmoles of dATP, 22 nmoles of dCTP, 19 nmoles of dGTP and 10 nmoles of TTP.

The mixture was heated to 85° C. for five minutes and allowed to cool to ambient temperature. Five units of the Klenow fragment of *E. coli* DNA polymerase I were added and the temperature was maintained for 15 minutes. After that time, the mixture was again heated to 85° C. for five minutes and allowed to cool. Five units of the Klenow fragment were again added and the reaction was carried out for 15 minutes. The heating, cooling and synthesis steps were repeated eleven more times.

After the final repetition, a 5μl aliquot was removed from the reaction mixture. This was heated to 85° C. for three minutes and allowed to cool to ambient temperature. 12.5 pmoles of α-P^{32}-deoxycytidine triphosphate and 5 units of Klenow fragment were added and the reaction was allowed to proceed for 15 minutes. The labeled products were examined by polyacrylamide gel electrophoresis. The FokI digest was labeled in a similar fashion and served as a control and molecular weight markers. The only heavily labeled band visible after the 13 cycles was the intended 25 base pair sequence.

EXAMPLE 2

The desired sequence to be amplified was a 94 base pair sequence contained within the human beta-globin gene and spanning the MstII site involved in sickle cell anemia. The sequence has the nucleotide sequence shown in FIG. 1.

I. Synthesis of Primers

The following two oligodeoxyribonucleotide primers were prepared by the method described below:

5′ CACAGGGCAGTAACG 3′ Primer A

and

5′ TTTGCTTCTGACACA 3′ Primer B

Automated Synthesis Procedures: The diethylphosphoramidites, synthesized according to Beaucage and Caruthers (*Tetrahedron Letters* (1981) 22:1859–1862), were sequentially condensed to a nucleoside derivatized controlled pore glass support using a Biosearch SAM-1. The procedure included detritylation with trichloroacetic acid in dichloromethane, condensation using benzotriazole as activating proton donor, and capping with acetic anhydride and dimethylaminopyridine in tetrahydrofuran and pyridine. Cycle time was approximately 30 minutes. Yields at each step were essentially quantitative and were determined by collection and spectroscopic examination of the dimethoxytrityl alcohol released during detritylation.

Oligodeoxyribonucleotide Deprotection and Purification Procedures: The solid support was removed from the column and exposed to 1 ml concentrated ammonium hydroxide at room temperature for four hours in a closed tube. The support was then removed by filtration and the solution containing the partially protected oligodeoxyribonucleotide was brought to 55° C. for five hours. Ammonia was removed and the residue was applied to a preparative polyacrylamide gel. Electrophoresis was carried out at 30 volts/cm for 90 minutes after which the band containing the product was identified by UV shadowing of a fluorescent plate. The band was excised and eluted with 1 ml distilled water overnight at 4° C. This solution was applied to an Altech RP18 column and eluted with a 7–13% gradient of acetonitrile in 1% ammonium acetate buffer at pH 6.0. The elution was monitored by UV absorbance at 260 nm and the appropriate fraction collected, quantitated by UV absorbance in a fixed volume and evaporated to dryness at room temperature in a vacuum centrifuge.

Characterization of Oligodeoxyribonucleotides: Test aliquots of the purified oligonucleotides were ^{32}P la-

beled with polynucleotide kinase and γ-^{32}P-ATP. The labeled compounds were examined by autoradiography of 14–20% polyacrylamide gels after electrophoresis for 45 minutes at 50 volts/cm. The procedure verifies the molecular weight. Base composition was determined by digestion of the oligodeoxyribonucleotide to nucleosides by use of venom diesterase and bacterial alkaline phosphatase and subsequent separation and quantitation of the derived nucleosides using a reverse phase HPLC column and a 10% acetonitrile, 1% ammonium acetate mobile phase.

II. Source of DNA

A. Extraction of Whole Human Wild-Type DNA

Human genomic DNA homozygous for normal β-globin was extracted from the cell line Molt4 (obtained from Human Genetic Mutant Cell Repository and identified as GM2219c) using the technique described by Stetler et al., *Proc. Nat. Acad. Sci. USA* (1982), 79:5966–5970.

B. Construction of Cloned Globin Genes

A 1.9 kb BamHI fragment of the normal β-globin gene was isolated from the cosmid pFC11 and inserted into the BamHI site of pBR328 (Soberon, et al., Gene (1980) 9:287–305). This fragment, which encompasses the region that hybridizes to the synthetic 40-mer probe, includes the first and second exons, first intron, and 5′ flanking sequences of the gene (Lawn et al., Cell (1978), 15:1157–1174). This clone was designated pBR328:HbAand deposited under ATCC No. 39,698 on May 25, 1984.

The corresponding 1.9 kb BamHI fragment of the sickle cell allele of β-globin was isolated from the cosmid pFC12 and cloned as described above. This clone was designated pBR328:HbS and deposited under ATCC No. 39,699 on May 25, 1984.

Each recombinant plasmid was transformed into and propagated in *E. coli* MM294 (ATCC No. 39,607).

C. Digestion of Cloned Globin Genes with MstII

A total of 100 μg each of pBR328:HbA and pBR328:HbS were individually digested with 20 units of MstII (New England Biolabs) for 16 hours at 37° C. in 200 μl of 150 mM NaCl, 12 mM Tris HCl (pH 7.5), 12 mM $MgCl_2$, 1 mM dithiothreitol (DTT), and 100 μg/ml bovine serum albumin (BSA). The products are designated pBR328:HbA/MstII and pBR328:HbS/MstII, respectively.

III. Polymerase Chain Reaction

To 100 μl of buffer consisting of 60 mM sodium acetate, 30 mM Tris acetate and 10 mM magnesium acetate at pH 8.0 was added 2 μl of a solution containing 100 picomoles of Primer A (of the sequence d(CACAGGGCACTAACG)), 100 picomoles of Primer B (of the sequence d(TTTBCTTCTGACACA)) and 1000 picomoles each of dATP, dCTP, dGTP and TTP. In addition, one of the following sources of DNA described above was added:

10 μg whole human wild-type DNA (Reaction I)
0.1 picomole pBR328:HbA (Reaction II)
0.1 picomole pBR328:HbS (Reaction III)
0.1 picomole pBR328:HbA/MstII (Reaction IV)
0.1 picomole pBR328:HbS/MstII (Reaction V)
No target DNA (Reaction VI)

Each resulting solution was heated to 100° C. for four minutes and allowed to cool to room temperature for two minutes, whereupon 1 μl containing four units of Klenow fragment of *E. coli* DNA polymerase was added. Each reaction was allowed to proceed for 10 minutes, after which the cycle of adding the primers, nucleotides and DNA, heating, cooling, adding polymerase, and reacting was repeated nineteen times for Reaction I and four times for Reactions II–VI.

Four microliter aliquots of Reactions I and II removed before the first cycle and after the last cycle of each reaction were applied to a 12% polyacrylamide gel 0.089M in Tris-borate buffer at pH 8.3 and 2.5 mM in EDTA. The gel was electrophoresed at 25 volts/cm for four hours, transferred to a nylon membrane serving as solid phase support and probed with a 5′-^{32}P-labeled 40 bp synthetic fragment, prepared by standard techniques, of the sequence

5′d(TCCTGAGGAGAAGTCTGCCGT-
TACTGCCCTGTGGGGCAAG)3′

in 30% formamide, 3×SSPE, 5×Denhardt's, 5% sodium dodecyl sulfate at pH 7.4. FIG. 2 is an autoradiograph of the probed nylon membrane for Reactions I and II. Lane 1 is 0.1 picomole of a 58-bp synthetic fragment control one strand of which is complementary to the above probe. Lane 2 is 4 μl of Reaction I prior to the first amplification cycle. Lane 3 is 4 μl of Reaction I after the 20th amplification cycle. Lane 4 is 4 μl of Reaction II after five amplification cycles. Lane 5 is a molecular weight standard consisting of a FokI (New England Biolabs) digest of pBR322 (New England Biolabs) labeled with alpha-^{32}P-dNTPs and polymerase. Lane 3 shows that after twenty cycles the reaction mixture I contained a large amount of the specific sequence of the proper molecular weight and no other detectable products. Reaction mixture II after five cycles also contained this product, as well as the starting nucleic acid and other products, as shown by Lane 4.

To 5.0 μl aliquots of Reactions II–VI after the fourth cycle were added 5 pmoles of each primer described above. The solutions were heated to 100° C. for four minutes and allowed to equilibrate to room temperature. Three pmoles each of alpha-^{32}P-dATP, alpha-^{32}P-dCTP, alpha-^{32}P-dGTP and alpha-^{32}P-TTP and four units of Klenow fragment were added. The reaction, in a final volume of 10 μl and at the salt concentrations given above, was allowed to proceed for 10 minutes. The polymerase activity was terminated by heating for 20 minutes at 60° C. Four μl aliquots of Reactions II–VI were loaded onto a 12% polyacrylamide gel 0.089M in Tris-borate buffer at pH 8.3 and 2.5 mM in EDTA. The gel was electrophoresed at 25 volts/cm for four hours after which autoradiography was performed.

FIG. 3 is an autoradiograph of the electrophoresis. Lane 1 is a molecular weight standard, Lane 2 is Reaction II, Lane 3 is Reaction III, Lane 4 is Reaction IV and Lane 5 is Reaction V. Another lane for Reaction VI with no DNA as control had no images in any of the lanes. It can be seen from the figure that the 94--bp fragment predicted from the target DNA was present only where intact β-globin DNA sequences were available for amplification, i.e., pBR328: HbA (Lane 2), pBR328: HbS (Lane 3) and pBR328: HbS/MstII (Lane 5). MstII digestion cuts pBR328: HbA in the 94-mer sequence rendering it incapable of being amplified, and the 94-mer band does not appear in Lane 4. In contrast, the 94-mer sequence in pBR328: HbS does not cut when the plasmid is digested with MstII and thus is available for amplification as shown in Lane 5.

FIG. 4 illustrates the chain reaction for three cycles in amplifying the 94-bp sequence. PC01 and PC02 are

Primers A and B. The numbers on the right indicate the cycles, whereas the numbers on the left indiciates the cycle number in which a particular molecule was produced.

EXAMPLE 3

This example illustrates amplification of a 110 bp sequence spanning the allelic MstII site in the human hemoglobin gene.

A total of 1.0 microgram whole human DNA, 100 picomoles D(ACACAACTGTGTTCACTAGC) and 100 picomoles (dCAACTTCATCCACGTTCACC), the primers having been prepared by the technique of Example 2, were dissolved in 100 μl of a solution which was:

1.5 mM in each of the four deoxyribonucleoside triphosphates
30 mM in Tris acetate buffer at pH 7.9
60 mM in sodium acetate
10 mM in magnesium acetate
0.25 mM in dithiothreitol

The solution was heated to 100° C. for one minute and brought rapidly to 25° C. for one minute, after which was added 2.5 units Klenow fragment of DNA polymerase. The polymerase reaction was allowed to proceed for two minutes at 25° C., after which the cycle of heating, cooling, adding Klenow, and reacting was repeated as often as desired.

With a 70% efficiency at each cycle, 15 cycles resulted in the synthesis of 1.4 femtomoles of the desired 110 bp fragment of the β-globin gene.

EXAMPLE 4

This example illustrates amplification of a 240 bp sequence spanning the allelic MstII site in the human hemoglobin gene. This sequence contains NcoI, HinfI and MstII restriction sites.

To 100 μl of a mixture of 60 mM sodium acetate, 30 mM Tris acetate and 10 mM magnesium acetate at pH 8.0 containing 0.1 pmole pBR328: HbA was added 2 μof Solution A containing:

100 pmoles d(GGTTGGCCAATCTACTCCAGG) primer
100 pmoles d(TAACCTTGATACCAACCTGCCC) primer
1000 pmoles each of dATP, dCTP, dGTP and TTP

The two primers were prepared by the technique described in Example 2. The solution was heated to 100° C. for four minutes and allowed to cool in ambient air for two minutes, after which was added 1 μl containing four units Klenow fragment of *E. coli* DNA polymerase. The reaction was allowed to proceed for 10 minutes after which the cycle of solution A addition, heating, cooling, adding polymerase, and reacting was repeated three times. To a 5.0 μl aliquot of the reactions was added 5 picomoles of each oligonucleotide primer described above. The solution was heated to 100° C. for four minutes and allowed to come to ambient temperature, after which 3 picomoles each of the alpha-^{32}P-labeled deoxyribonucleoside triphosphates and 4 units Klenow fragment were added. The reaction, in a final volume of 10 μl and at the salt concentrations given above, was allowed to proceed for 10 minutes. The polymerase activity was terminated by heating for 20 minutes at 60° C. Two μl aliquots were digested with NcoI, MstII, or HinfI and loaded onto a 12% polyacrylamide gel 0.089M in Tris-borate buffer at pH 8.3 and 2.5 mM in EDTA. The gel was electrophoresed at 25 volts/cm for four hours and autoradiography was performed. FIG. 5 illustrates the autoradiograph of the electrophoresis, were Lane 1 is the molecular weight standard, Lane 2 is without digestion with enzyme (240 bp intact), Lane 3 is digestion with NcoI (131 and 109 bp), Lane 4 is digestion with MstII (149 and 91 bp), and Lane 5 is digestion with HinfI (144 and 96 bp). The autoradiograph is consistent with the amplification of the 240 bp sequence.

EXAMPLE 5

This example illustrates use of the process herein to detect sickle cell anemia by sequential digestion.

Synthesis and Phosphorylation of Oligodeoxyribonucleotides

A labeled DNA probe, RS06, of the sequence

5′ *CTGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGG 3′

where * indicates the label, and an unlabeled blocking oligomer, RS10, of the sequence:

3′ GACAGAGGTCACCTCTTCAGACGGCAATGACGGGACACCC 5′

which has three pair mismatches with RS06 were synthesized according to the procedures provided in Example 2(I). The probe RS06 was labeled by contacting five pmole thereof with 4 units of T4 polynucleotide kinase (New England Biolabs) and 50 pmole γ-^{32}P-ATP (New England Nuclear, about 7200 Ci/mmole) in a 40 μl reaction volume containing 70 mM Tris buffer (pH 7.6), 10 mM $MgCl_2$, 1.5 mM spermine, and 2.5 mM dithiothreitol for 90 minutes at 37° C. The total volume was then adjusted to 100 μl with 25 mM EDTA and purified according to the procedure of Maniatis et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory, 1982), pp. 464–465 over a 1 ml Bio Gel P-4 spin dialysis column from Bio-Rad equilibrated with Tris-EDTA (TE) buffer (10 mM Tris buffer, 0.1 mM EDTA, pH 8.0). The labeled probe was further purified by electrophoresis on a 18% polyacrylamide gel (19:1 acrylamide:BIS, Bio-Rad) in Tris-boric acid-EDTA (TBE) buffer (89 mM Tris, 89 mM boric acid, 2.5 mM EDTA, pH 8.3) for 500 vhr. After localization by autoradiography, the portion of the gel containing the labeled probe was excised, crushed and eluted into 0.2 ml TE buffer overnight at 4° C. TCA precipitation of the reaction product indicated that the specific activity was 4.9 Ci/mmole and the final concentration was 20 pmole/ml.

The unlabeled RS10 blocking oligomer was used at a concentration of 200 pmole/ml.

Isolation of Human Genomic DNA from Cell Lines

High molecular weight genomic DNA was isolated from the lymphoid cell lines Molt4, SC-1 and GM2064 using essentially the method of Stetler et al., Proc. Natl. Acad. SC1. USA (1982), 79, 5966–5970 (for Molt4) and Maniatis et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory, 1982), pp. 280–281

Molt4 (Human Mutant Cell Repository, GM2219C) is a T cell line homozygous for normal β-globin, and SC-1, deposited with ATCC on Mar. 19, 1985, is an EBV-transformed B cell line homozygous for the sickle cell allele. GM2064 (Human Mutant Cell Repository, GM2064) was originally isolated from an individual homozygous for hereditary persistance of fetal hemoglobin (HPFH) and contains no beta- or delta- globin

Case 3:11-cv-00056-IEG -WVG Document 1-2 Filed 01/10/11 Page 82 of 90

gene sequences. All cell lines were maintained in RPMI-1640 with 10% fetal calf serum.

Isolation of Human Genomic DNA from Clinical Blood Samples

A clinical blood sample designated CH12 from a known sickle cell carrier (AS) was obtained from Dr. Bertram Lubin of Children's Hospital in Oakland, Calif. Genomic DNA was prepared from the buffy coat fraction, which is composed primarily of peripheral blood lymphocytes, using a modification of the procedure described by Nunberg et al., Proc. Nat. Acad. Sci. U.S.A., 75, 5553–5556 (1978).

The cells were resuspended in 5 ml Tris-EDTA-NaCl (TEN) buffer (10 mM Tris buffer pH 8, 1 mM EDTA, 10 mM NaCl) and adjusted to 0.2 mg/ml proteinase K, 0.5% SDS, and incubated overnight at 37° C. Sodium perchlorate was then added to 0.7 M and the lysate gently shaken for 1–2 hours at room temperature. The lysate was extracted with 30 ml phenol/chloroform (1:1), then with 30 ml chloroform, and followed by ethanol precipitation of the nucleic acids. The pellet was resuspended in 2 ml of TE buffer and RNase A added to 0.005 mg/ml. After digestion for one hour at 37° C., the DNA was extracted once each with equal volumes of phenol, phenol/chloroform, and chloroform, and ethanol precipitated. The DNA was resuspended in 0.5 ml TE buffer and the concentration was determined by absorbance at 260 nm.

Polymerase Chain Reaction to Amplify Selectively β-Globin Sequences

Two micrograms of genomic DNA was amplified in an initial 100 μl reaction volume containing 10 mM Tris buffer (ph 7.5), 50 mM NaCl, 10 mM $MgCl_2$, 150 pmole of Primer A of the sequence d(CACAGGGCACTAACG), and 150 pmole of Primer B of the sequence d(CTTTGCTTCTGACACA) and overlayed with about 100 μl mineral oil to prevent evaporation.

Each DNA sample underwent 15 cycles of amplification where one cycle is composed of three steps:

(1) Denature in a heat block set at 95° C. for two minutes.

(2) Transfer immediately to a heat block set at 30° C. for two minutes to allow primers and genomic DNA to anneal.

(3) Add 2 μl of a solution containing 5 units of the Klenow fragment of E. coli DNA polymerase I (New England Biolabs), 1 nmole each of dATP, dCTP, dGTP and TTP, in a buffer composed of 10 mM Tris (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$, and 4 mM dithiothreitol. This extension reaction was allowed to proceed for 10 minutes at 30° C.

After the final cycle, the reaction was terminated by heating at 95° C. for two minutes. The mineral oil was extracted with 0.2 ml of chloroform and discarded. The final reaction volume was 130 μl.

Hybridization/Digestion of Amplified Genomic DNA with Probes and DdeI/HinfI

Forty-five microliters of the amplified genomic DNA was ethanol precipitated and resuspended in an equal volume of TE buffer. Ten microliters (containing the pre-amplification equivalent of 154 ng of genomic DNA) was dispensed into a 1.5 ml Microfuge tube and 20 μl of TE buffer to a final volume of 30 μl. The sample was overlayed with mineral oil and denatured at 95° C. for 10 minutes. Ten microliters of 0.6 M NaCl containing 0.02 pmole of labeled RS06 probe was added to the tube, mixed gently, and immediately transferred to a 56° C. heat block for one hour. Four microliters of unlabeled RS10 blocking oligomer (0.8 pmole) was added and the hybridization continued for an additional 10 minutes at the same temperature. Five microliters of 60 mM $MgCl_2$/0.1% BSA and 1 μl of DdeI (10 units, New England Biolabs) were added and the reannealed DNA was digested for 30 minutes at 56° C. One microliter of HinfI (10 units, New England Biolabs) was then added and incubated for another 30 minutes. The reaction was stopped by the addition of 4 μl 75 mM EDTA and 6 μl tracking dye to a final volume of 61 μl.

The mineral oil was extracted with 0.2 ml chloroform, and 18 μl of the reaction mixture (45 ng genomic DNA) was loaded onto a 30% polyacrylamide mini-gel (19:1, Bio-Rad) in a Hoeffer SE200 apparatus. The gel was electrophoresed at approximately 300 volts for one hour until the bromphenol blue dye front migrated to 3.0 cm offorigin. The top of 1.5 cm of the gel was removed and the remaining gel was exposed for four days with one intensification screen at −70° C.

Discussion of Autoradiograph (FIG. 9)

Each lane contains 45 ng of amplified genomic DNA. Lane A contains Molt4 DNA; Lane B, CH12; Lane C, SC-1; and Lane D, GM2064. Molt4 represents the genotype of a normal individual with two copies of the β^A gene per cell (AA), CH12 is a clinical sample from a sickle cell carrier with one β^A and one β^S gene per cell (AS), and SC-1 represents the genotype of a sickle cell individual with two copies of the β^S gene per cell (SS). GM2064, which contains no beta- or delta-globin sequences, is present as a negative control.

As seen in the autoradiogram, the DdeI-cleaved, β^A-specific octamer is present only in those DNA's containing the β^A gene (Lanes A and B), and the HinfI-cleaved, β^S-specific trimer is present only in those DNA's containing the β^S gene (Lanes B and C). The presence of both trimer and octamer (Lane B) is diagnostic for a sickle cell carrier and is distinguishable from a normal individual (Lane A) with only octamer and a sickle cell afflicted individual (Lane C) with only trimer.

As a comparison, repeating the experiment described above using non-amplified genomic DNA revealed that the amplification increased the sensitivity of detection by at least 1000 fold.

EXAMPLE 6

This example illustrates direct detection of a totally unpurified single copy gene in whole human DNA on gels without the need for a labeled probe.

Using the technique described in Example 3, a 110-bp fragment from a sequence in the first exon of the beta-globin gene was amplified from 10 micrograms of whole human DNA after 20 cycles. This 110-bp fragment produced after 20 cylces was easily visualized on gels stained with ethidium bromide.

The sequence was not amplified when it was first cut with the restriction enzyme DdeI unless, as in the beta-globin S allele, the sequence does not contain the restriction site recognized by the enzyme.

EXAMPLE 7

A. A total of 100 fmoles pBR328 containing a 1.9 kb insert from the human beta-globin A allele, 50 nmoles each alpha-32P-dNTP at 500 Ci/mole, and 1 nmole of

each of the primers used in Example 3 were dissolved in a solution containing 100 μl 30 mM Tris-acetate at pH 7.9, 60 mM sodium acetate, 100 mM dithiothreitol, and 10 mM magnesium acetate. This solution was brought to 100° C. for two minutes and cooled to 25° C. for one minute. A total of 1 μl containing 4.5 units Klenow fregment of E. coli DNA polymerase I and 0.09 units inorganic pyrophosphatase was added to prevent the possible build-up of pyrophosphate in the reaction mixture, and the reaction was allowed to proceed for two minutes at 25° C., after which the cycle of heating, cooling, adding enzyme, and reacting was repeated nine times. Ten-μl aliquots were removed and added to 1 μl 600 mM EDTA after each synthesis cycle. Each was analyzed on a 14% polyacrylamide gel in 90 mM Tris-borate and 2.5 mM EDTA at pH 8.3 and 24 volts/cm for 2.5 hours. The completed gel was soaked for 20 minutes in the same buffer with the addition of 0.5 μg/ml ethidium bromide, washed with the original buffer, and photographed in UV light using a red filter.

The 110-bp fragment produced was excised from the gel under ultraviolet light and the incorporated ^{32}p counted by Cerenkov radiation. An attempt to fit the data to an equation of the form: pmoles/10 μl=0.01 $[(1+y)^N-yN-1]$, where N represents the number of cycles and y the fractional yield per cycle, was optimal with y=0.619. This indicates that a significant amplification is occurring.

B. The above experiment was repeated except that 100 nmoles of each dNTP was added to a 100 μl reaction, no radiolabel was employed, and aliquots were not removed at each cycle. After 10 cycles the reaction was terminated by boiling for two minutes and rehybridization was performed at 57° C. for one hour. The sequence of the 110-bp product was confirmed by subjecting 8 μl aliquots to restriction analysis by addition of 1 μl bovine serum albumin (25 mg/ml) and 1 μl of the appropriate restriction enzyme (HinfI, MnlI, MstII, NcoI) and by reaction at 37° C. for 15 hours. PAGE was performed as described above.

EXAMPLE 8

This example illustrates the use of different primers to amplify various fragments of pBR328 and 322.

A. The experiment described in Example 7A was repeated except using the following primers: d(TTTGCTTCTGACACAACTGTGTTCACTAGC) and d(GCCTCACCACCAACTTCATCCACGTTCACC) to produce a 130-bp fragment of pBR328.

B. The experiment described in Example 7A was repeated except using the following primers: d(GGTTGGCCAATCTACTCCCAGG) and d(TGGTCTCCTTAAACCTGTCTTG) to produce a 262-bp fragment of pBR328. The reaction time was 20 minutes per cycle.

C. The experiment described in Example 8B was repeated except that 100 fmoles of an MstII digest of pBR328 containing a 1.9 kb insert from the human beta-globin S allele was used as initial template. This plasmid was cleaved several times by MstII but not inside the sequence to be amplified. In addition, the primers employed were as follows:

d(GGTTGGCCAATCTACTCCCAGG) and
d(TAACCTTGATACCAACCTGCCC)

to produce a 240-bp fragment.

D. The experiment described in Example 7B was repeated except that 100 fmoles of an NruI digest of pBR322 was used as template, 200 nmoles of each dNTP were used in the 100 μl reaction, and the primers were:

d(TAGGCGTATCACGAGGCCCT) and
d(CTTCCCCATCGGTGATGTCG) to produce a 500-bp fragment from pBR322. Reaction times were 20 minutes per cycle at 37° C. Final rehybridization was 15 hours at 57° C. Electrophoresis was on a 4% agarose gel.

EXAMPLE 9

This example illustrates the invention process wherein an in vitro mutation is introduced into the amplified segment.

A. A total of 100 fmoles of pBR322 linearized with NruI, 1 nmole each of the primers:

d(CGCATTAAAGCTTATCGATG) and
d(TAGGCGTATCACGAGGCCCT)

designed to produce a 75-bp fragment, 100 nmole each dNTP, in 100 μl 40 mM Tris at pH 8, 20 mM in $MgCl_2$, 5 mM in dithiothreitol, and 5 mg/ml bovine serum albumin were combined. The mixture was brought to 100° C. for one minute, cooled for 0.5 minutes in a water bath at 23° C., whereupon 4.5 units Klenow fragment and 0.09 units inorganic pyrophosphatase were added, and a reaction was allowed to proceed for three minutes. The cycle of heating, cooling, adding enzymes, and reacting was repeated nine times. The tenth reaction cycle was terminated by freezing and an 8-μl aliquot of the reaction mixture was applied to a 4% agarose gel visualized with ethidium bromide.

B. The experiment described in Example 9A was repeated except that the oligonucleotide primers employed were:

d(CGCATTAAAGCTTATCGATG) and
d(AATTAATACGACTCACTATAAGGGAGATAGGCGTATCACGAGGCCCT).

These primers are designed to produce a 101-bp fragment, 26 nucleotides of which (in the second listed primer) are not present in pBR322. These nucleotides represent the sequence of the T7 promoter, which was appended to the 75-bp sequence from pBR322 by using the primer with 20 complementary bases and a 26-base 5′ extension. The procedure required less than two hours and produced two picomoles of the relatively pure 101-bp fragment from 100 fmoles of pBR322.

The T7 promoter can be used to initiate RNA transcription. T7 polymerase may be added to the 101-bp fragment to produce singlestranded RNA.

C. The experiment described in Example 8D was repeated except that the oligonucleotide primers employed were as follows:

d(TAGGCGTATCACGAGGCCCT) and
d(CCAGCAAGACGTAGCCCAGC)

to produce a 1000-bp fragment from pBR322.

D. The experiment described in Example 9C was repeated except that the oligonucleotide primers employed were as follows:

d(TAGGCGTATCACGAGGCCCT) and
d(AATTAATACGACTCACTATAGGGAGATAGGCGTATCACGAGGCCCT)

so as to produce a 1026-bp fragment, 26 nucleotides of which (in the second listed primer) are not present in pBR322 and represent the T7 promoter described above. The promoter has been inserted adjacent to a 1000-bp fragment from pBR322.

The results indicate that a primer which is not a perfect match to the template sequence but which is none-

theless able to hybridize sufficiently to be enzymatically extended produces a long product which contains the sequence of the primer rather than the corresponding sequence of the original template. The long product serves as a template for the second primer to introduce an in vitro mutation. In further cycles this mutation is amplified with an undiminished efficiency, because no further mispaired primings are required. In this case, a primer which carries a non-complementary extension on its 5′ end was used to insert a new sequence in the product adjacent to the template sequence being copied.

E. Because the reaction with polymerase generates pyrophosphate and is theoretically reversible (Kornberg, A., DNA Replication, W. H. Freeman, San Francisco, 1980), the effect of including an inorganic pyrophosphatase to avoid potential pyrophosphorolysis of the product was examined. Qualitative polyacrylamide gel electrophoresis examination of reactions plus and minus pyrophosphatase demonstrated a minor but significant increase in homogeneity of product as a result of the inclusion of this enzyme.

EXAMPLE 10

This example illustrates employing nested sets of primers to decrease the background in the amplification of single copy genes.

Whole human DNA homozygous for the wild-type betaglobin allele was subjected to twenty cycles of amplification as follows: A total of 10 μg DNA, 200 picomoles each of the primers:

d(ACACAACTGTGTTCACTAGC) and
d(CAACTTCATCCACGTTCACC)

and 100 nanomoles each dNTP in 100 μl of 30 mM Tris-acetate pH 7.9, 60 mM sodium acetate, 10 mM dithiothreitol, and 10 mM magnesium acetate were heated to 100° C. for one minute, cooled to 25° C. for one minute, and treated with 2 units Klenow fragment for two minutes. The cycle of heating, cooling and adding Klenow was repeated 19 times. A ten-μl aliquot was removed from the reaction mixture and subjected to a further ten cycles of amplification using each of the primers:

d(CAGACACCATGGTGCACCTGACTCCTG) and
d(CCCCACAGGGCAGTAACG-GCAGACTTCTCC),

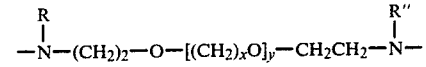
which amplify a 58-bp fragment contained within the 110-bp fragment produced above. This final ten cycles of amplification was accomplished by diluting the 10-μl aliquot into 90 μl of the fresh Tris-acetate buffer described above containing 100 nanomoles each dNTP and 200 pmoles of each primer. Reaction conditions were as above. After ten cycles a 10-μl aliquot (corresponding to 100 nanograms of the original DNA) was applied to a 6% NuSieve (FMC Corp.) agarose gel and visuallized using ethidium bromide.

FIG. **10** illustrates this gel illuminated with UV light and photographed through a red filter as is known in the art. Lane 1 is molecular weight markers. Lane 2 is an aliquot of the reaction described above. Lane 3 is an aliquot of a reaction identical to that described above, except that the original wild-type DNA was cleaved with DdeI prior to amplification. Lane 4 is an aliquot of a reaction identical to that described above, except that human DNA homozygous for the sickle betaglobin allele was treated with DdeI prior to amplification (the sickle allele does not contain a DdeI site in the fragment being amplified here). Lane 5 is an aliquot of a reaction identical to that described above, except that salmon sperm DNA was substituted for human DNA. Lane 6 is an aliquot of a reaction identical to that described above, except that the aliquot was treated with DdeI after amplification (DdeI should convert the 58-bp wild-type product into 27-and 31-bp fragments). Lane 7 is an aliquot of the Lane 4 material treated with DdeI after amplification (the 58-bp sickle product contains no DdeI site).

Detection of a 58-bp fragment representative of a singlecopy gene from one microgram of human DNA using only ethidium bromide staining of an agarose gel requires an amplification of about 500,000-fold. This was accomplished by using the two nested sets of oligonucleotide primers herein. The first set amplifies the 110-bp fragment and the inner nested set amplifies a sub-fragment of this product up to the level of convenient detection shown in FIG. **10**. This procedure of using primers amplifying a smaller sequence contained within the sequence being amplified in the previous amplification process and contained in the extension products of the other primers allows one to distinguish the wild-type from the sickle allele at the betaglobin locus without resorting to either radiosotopic or otherwise cumbersome methodology such as that of Conner et al., Proc. Natl. Acad. Sci. U.S.A., 80:278 (1983) and Leary et al., Proc. Natl. Acad. Sci. U.S.A., 80:4045 (1983).

EXAMPLE 11

The present process is expected to be useful in detecting, in a patient DNA sample, a specific sequence associated with an infectious disease such as, e.g., Chlamydia using a biotinylated hybridization probe spanning the desired amplified sequence and using the process described in U.S. Pat. No. 4,358,535, supra. The biotinylated hybridization probe may be prepared by intercalation and irradiation of a partially double-stranded DNA with a 4′-methylene substituted 4,5′-8-trimethylpsoralen attached to biotin via a spacer arm of the formula:



where R is —H or a —CHO group, R″ is —H, x is a number from 1 to 4, and y is a number from 2 to 4, as described in U.S. Pat. No. 4,582,789 issued Apr. 15, 1986 to K. Mullis et al., the disclosure of which is incorporated herein by reference. Detection of the biotinyl groups on the probe may be accomplished using a streptavidin-acid phosphatase complex commercially obtainable from Enzo Biochemical using the detection procedures suggested by the manufacturer in its brochure. The hybridized probe is seen as a spot of precipitated stain due to the binding of the detection complex, and the subsequent reaction catalyzed by acid phosphatase, which produces a precipitable dye.

Deposit of Materials

The cell line SC-1 (CTCC #0082) was deposited on Mar. 19, 1985 with the American Type Culture Collection (ATCC), 12301 Parklawn Drive, Rockville, Md. 20852 USA, with ATCC Accession No. CRL#8756, The deposit of SC-1 was made pursuant to a contract between the ATCC and the assignee of this patent application, Cetus Corporation. The contract with ATCC

provides for permanent availability of the progeny of this cell line to the public on the issuance of the U.S. patent describing and identifying the deposit or the publications or upon the laying open to the public of any U.S. or foreign patent application, whichever comes first, and for availability of the progeny of this cell line to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 CFR §122 and the Commissioner's rules pursuant thereto (including 37 CFR §1.14 with particular reference to 886 OG 638). The assignee of the present application has agreed that if the cell line on deposit should die or be lost or destroyed when cultivated under suitable conditions, it will be promptly replaced on notification with a viable culture of the same cell line.

In summary, the present invention is seen to provide a process for amplifying one or more specific nucleic acid sequences using a chain reaction in which primer extension products are produced which can subsequently act as templates for further primer extension reactions. The process is especially useful in detecting nucleic acid sequences which are initially present in only very small amounts.

Other modifications of the above described embodiments of the invention which are obvious to those of skill in the area of molecular biology and related disciplines are intended to be within the scope of the following claims.

What is claimed is:

1. A process for amplifying at least one specific nucleic acid sequence contained in a nucleic acid or a mixture of nucleic acids wherein each nucleic acid consists of two separate complementary strands, of equal or unequal length, which process comprises:

(a) treating the strands with two oligonucleotide primers, for each different specific sequence being amplified, under conditions such that for each different sequence being amplified an extension product of each primer is synthesized which is complementary to each nucleic acid strand, wherein said primers are selected so as to be sufficiently complementary to different strands of each specific sequence to hybridize therewith such that the extension product synthesized from one primer, when it is separated from its complement, can serve as a template for synthesis of the extension product of the other primer;

(b) separating the primer extension products from the templates on which they were synthesized to produce single-stranded molecules; and

(c) treating the single-stranded molecules generated from step (b) with the primers of step (a) under conditions that a primer extension product is synthesized using each of the single strands produced in step (b) as a template.

2. The process of claim **1**, wherein steps (b) and (c) are repeated at least once.

3. The process of claim **1**, wherein said step (b) is accomplished by denaturing.

4. The process of claim **3**, wherein said denaturing is caused by heating.

5. The process of claim **1**, wherein said step (b) is accomplished using the enzyme helicase.

6. The process of claim **1**, wherein steps (a) and (c) are accomplished using an enzyme.

7. The process of claim **6**, wherein said enzyme is selected from the group consisting of *E. coli* DNA polymerase I, Klenow fragment of *E. coli* DNA polymerase I, T4 DNA polymerase, reverse transcriptase wherein the template is RNA on DNA and the extension product is DNA, and an enzyme that after being exposed to a temperature of about 65°–90° C. forms said extension products at the temperature of reaction during steps (a) and (c).

8. The process of claim **7**, wherein said nucleic acid is double stranded and its strands are separated by denaturing before or during step (a).

9. The process of claim **1**, wherein said nucleic acid is DNA and said primers are oligodeoxyribonucleotides.

10. The process of claim **1**, wherein said nucleic acid is messenger RNA.

11. The process of claim **1** wherein said mixture of nucleic acids used in step (a) is the product of step (c).

12. The process of claim **11**, wherein the primers employed are different from the primers employed in the process for producing the product of step (c) used in step (a).

13. The process of claim **12**, wherein the primers employed result in the amplification of a smaller sequence contained within the sequence being amplified in the process for producing the product of step (c) used in step (a).

14. The process of claim **1**, wherein said steps are carried out simultaneously above room temperature using an enzyme that after exposed to a temperature of about 65°–90° C. forms said extension products at the temperature of reaction during steps (a) and (c).

15. The process of claim **1**, wherein the two primers in steps (a) and (c) are each present in a molar ratio of at least 1000:1 primer:complementary strand.

16. The process of claim **1**, wherein the nucleic acid sequence(s) to be modified is contained in a mixture of nucleic acids resulting from a chemical synthesis.

17. The process of claim **1**, wherein at least one primer contains at least one nucleotide which is not complementary to the specific sequence to be amplified.

18. The process of claim **17** wherein one primer comprises an oligonucleotide with 20 complementary nucleotides and, at its 5′ end, a T7 promoter containing 26 noncomplementary nucleotides.

19. A process for amplifying a specific nucleic acid sequence contained in double-stranded DNA which process comprises:

(a) separating the strands of the DNA by physical, chemical or enzymatic means;

(b) treating the single strands with two oligodeoxyribonucleotide primers, in a molar excess of primer: its complementary strand, under conditions such that an extension product of each primer is synthesized, using *E. coli* DNA polymerase I or Klenow fragment thereof, which extension product is complementary to each DNA strand, wherein said primers are selected so as to be sufficiently complementary to different strands of each specific sequence to hybridize therewith such that the extension product synthesized from one primer, when it is separated from its complement, can serve as a template for synthesis of the extension product of the other primer;

(c) separating the primer extension products from the templates on which they are synthesized to produce single-stranded molecules by physical, chemical or enzymatic means; and

(d) treating the single-stranded molecules generated from step (c) with the two primers of step (b), in a

molar excess of primer: its complementary molecule, under conditions such that a primer extension product is synthesized, using *E. coli* DNA polymerase I or Klenow fragment thereof, and using each of the single strands produced in step (c) as a template.

20. The process of claim **19**, wherein steps (c) and (d) are repeated at least once.

21. The process of claim **1**, wherein, due to the degeneracy of the genetic code, a collection of primers is employed for each complementary strand, the sequence of one of which primers is exactly complementary to said complementary strand over the length of the primer.

* * * * *

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 4,683,202
DATED : July 28, 1987
INVENTOR(S) : Kary B. Mullis

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Title page,
After [*] Notice, please replace "The portion of the term of this patent subsequent to Jul. 28, 2004 has been disclaimed." with -- This patent is subject to a terminal disclaimer. --

Signed and Sealed this

First Day of July, 2003

JAMES E. ROGAN
Director of the United States Patent and Trademark Office

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

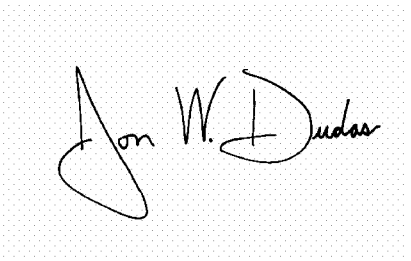
PATENT NO. : 4,683,202
APPLICATION NO. : 06/791308
DATED : July 28, 1987
INVENTOR(S) : Kary B. Mullis

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

TITLE PAGE, after "[*] Notice:" replace "The portion of the term of this patent subsequent to Jul. 28, 2004 has been disclaimed." with --This patent is subject to a terminal disclaimer.--

Signed and Sealed this

Fifth Day of June, 2007



JON W. DUDAS
Director of the United States Patent and Trademark Office

REEXAMINATION CERTIFICATE (1388th)

United States Patent [19]

Mullis

[11] **B1 4,683,202**

[45] Certificate Issued * **Nov. 27, 1990**

[54] **PROCESS FOR AMPLIFYING NUCLEIC ACID SEQUENCES**

[75] Inventor: **Kary B. Mullis**, Kensington, Calif.

[73] Assignee: **Cetus Corporation**

Reexamination Request:
No. 90/001,903, Dec. 6, 1989
No. 90/001,955, Mar. 9, 1990

Reexamination Certificate for:

Patent No.:	**4,683,202**
Issued:	**Jul. 28, 1987**
Appl. No.:	**791,308**
Filed:	**Oct. 25, 1985**

[*] Notice: The portion of the term of this patent subsequent to Jul. 28, 2004 has been disclaimed.

[51] **Int. Cl.**[5] **C12P 19/34**; C12N 15/00; C12N 1/00; C07H 21/04

[52] **U.S. Cl.** **435/91**; 435/91; 435/172.3; 435/317.1; 536/27; 536/28; 536/29; 935/17; 935/18; 935/16

[58] **Field of Search** 435/91, 172.3, 317.1; 536/27, 28, 29; 935/17, 18

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,555,486	11/1985	Bahl et al.	435/91
4,582,788	4/1986	Erlich	435/6
4,737,462	4/1988	Mark et al.	435/252.33
4,853,332	8/1989	Mark et al.	435/252.33

OTHER PUBLICATIONS

Khorana, Progress Report, Grant No. CA11981-02, 09-01-1971 through 08-31-1972.
Zoller et al; DNA 3: 479 (1984).
BioLogicals Advertisement; Science, Dec. 11, 1981.
New England Biolabs Catalog, 1983/84, p. 29.
Molecular Cloning A Laboratory Manual, 1982, Maniatis et al, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp. 113-132 and 230-238.
Kaledin et al; Biokhimiya 45(4), 644 (1980), published by Plenum Publishing Corporation.
Chien et al; J. Bacteriol. 127: 1550 (1976).
Anonymous; Nature 341: 570 (1989).
Kleppe et al; J. Mol. Biol. 56: 341 (1976).
Khorana et al; J. Mol. Biol. 72: 209 (1972).
Panet et al; J. Biol. Chem. 249: 5213 (1974).
Besmer et al., 1972, *J. Mol. Biol.* 72:503-522.

Primary Examiner—James Martinell

[57] **ABSTRACT**

The present invention is directed to a process for amplifying any desired specific nucleic acid sequence contained in a nucleic acid or mixture thereof. The process comprises treating separate complementary strands of the nucleic acid with a molar excess of two oligonucleotide primers, and extending the primers to form complementary primer extension products which acts as templates for synthesizing the desired nucleic acid sequence. The steps of the reaction may be carried out stepwise or simultaneously and can be repeated as often as desired.

REEXAMINATION CERTIFICATE ISSUED UNDER 35 U.S.C. 307

NO AMENDMENTS HAVE BEEN MADE TO THE PATENT

AS A RESULT OF REEXAMINATION, IT HAS BEEN DETERMINED THAT:

The patentability of claims 1–21 is confirmed.

* * * * *